Folate Metabolism and Biofortification in Plants

Dieter Blancquaert

Ghent University
Faculty of Sciences
Department of Physiology

FOLATE METABOLISM AND BIOFORTIFICATION IN PLANTS

Dieter Blancquaert

Promoter: Prof. Dr. Dominique Van Der Straeten
Laboratory of Functional Plant Biology
Ledeganckstraat 35
B-9000 Ghent, Belgium

This thesis is submitted in partial fulfillment of the requirements for the degree of PhD in Sciences, Biotechnology

This work was conducted at the Laboratory of Functional Plant Biology.

This work was supported by a PhD grant from the Fund for Scientific Research Flanders (FWO).

JURY MEMBERS

Prof. Dr. Dominique Van Der Straeten (Promoter)
Faculty of Sciences, Department of Physiology, Ghent University

Prof. Dr. Wout Boerjan (Chair)
Faculty of Sciences, Department of Plant Systems Biology, VIB/Ghent University

Prof. Dr. Apr. Christophe Stove (Reading Commission)
Faculty of Pharmaceutical Sciences, Department of Bioanalysis, Ghent University

Dr. Fabrice Rébeillé (Reading Commission)
Laboratoire de Physiologie Cellulaire Végétale, Institut de Recherches en Technologies et Sciences pour le vivant, Université Joseph Fourier Grenoble I, France

Prof. Dr. Han Asard (Reading Commission)
Onderzoeksgroep Moleculaire Plantenfysiologie en Biotechnologie, Universiteit Antwerpen

Dr. Pieter Ouwerkerk (Reading Commission)
Bayer Cropscience N.V., Zwijnaarde

Prof. Dr. Apr. Willy Lambert
Faculty of Pharmaceutical Sciences, Department of Bioanalysis, Ghent University

Prof. Dr. Bart Braeckman
Faculty of Sciences, Department of Biology, Ghent University

TABLE OF CONTENTS

Abbreviations

10-FDF, 10-formylTHF deformylase

10-FS, 10-formylTHF synthase

ADC, aminodeoxychorismate

ADCL, aminodeoxychorismate lyase

ADCS, aminodeoxychorismate synthase

C1, carbon-one

DALY, Disability-Adjusted Life Years

DHC, 5,10-methylene THF dehydrogenase/5,10-methenylTHF cyclohydrolase

DHF, dihydrofolate

DHFR, dihydrofolate reductase

DHFS, dihydrofolate synthetase

DHNA, dihydroneopterin aldolase

DHP, dihydropteroate

DHPS, dihydropteroate synthase

DPA, days post-anthesis

dTMP, deoxythymidine monophosphate

dUTP, deoxyuridine triphosphate

FA, folic acid

FBP, folate binding protein

FPGS, folypolyglutamate synthetase

FW, fresh weight

GDC, glycine decarboxylation complex

Glb-1, globulin-1

GluB1, glutelin B1

GluB4, glutelin B4

GMO, Genetically Modified Organism

GRAS, Generally Regarded As Safe

GTP, guanosine triphosphate

GTPCHI, guanosine triphosphate cyclohydrolase 1

HMDHP, hydroxymethyldihydropterin

HPPK, dihydropterin pyrophosphokinase

MS, methionine synthase

MTases, methyltransferases

MTHFR, 5,10-methyleneTHF reductase

NTD, neural tube defect

p-ABA, para-aminobenzoate

RDA, Recommended Daily Allowance

RDI, Recommended/Reference Daily Intake

SAH, S-adenosylhomocysteine

SAM, S-adenosylmethionine

SHMT, serine hydroxymethyltransferase

THF, tetrahydrofolate

Introduction

Folates and folic acid: from fundamental research toward sustainable health

Folates and folic acid: from fundamental research toward sustainable health

Dieter Blancquaert,[§] Sergei Storozhenko,[§] Karen Loizeau,[†] Hans De Steur,[‡] Veerle De Brouwer,[¶] Jacques Viaene,[‡] Stéphane Ravanel,[†] Fabrice Rébeillé,[†] Willy Lambert,[¶] and Dominique Van Der Straeten[§]

[§] Unit Plant Hormone Signaling and Bio-imaging, Department of Physiology, Ghent University, K.L. Ledeganckstraat 35, 9000 Gent, Belgium

[†] Laboratoire de Physiologie Cellulaire Végétale, UMR5168 CNRS-CEA-INRA-Université Joseph Fourier Grenoble I, Institut de Recherches en Technologies et Sciences pour le Vivant, CEA-Grenoble, 17 rue des Martyrs, 38054 Grenoble Cedex 9, France

[‡] Division Agro-Food Marketing, Department of Agricultural Economics, Ghent University, Coupure Links 653, B-9000 Gent, Belgium

[¶]Laboratory of Toxicology, Ghent University, Harelbekestraat 72, B-9000 Gent, Belgium

ABSTRACT

Folates are of paramount importance in one-carbon metabolism of most organisms. Plants and micro-organisms are able to synthesize folates *de novo*, making them the main dietary source for humans and animals, which are dependent on food or feed supplies for folates. Folate deficiency is an increasing problem in the developing, as well as in the developed regions of the world, affecting millions of people. Different strategies, such as food fortification and folic acid supplementation, remain far from accessible for the poor rural populations in developing countries. Increasing knowledge concerning folate biosynthesis, transport and catabolism does not only deepen our insight on the regulation of folate metabolism but also provides the keys towards folate enhancement through metabolic engineering in bacteria, as well as in plants. Recently, promising results were obtained using such an approach, but further fundamental research is a prerequisite to develop a practicable solution to fight folate deficiency. In parallel, progress in the development and improvement of folate analysis has been made. Here, we provide the state-of-the-art of folate biosynthesis, catabolism, and salvage. Finally, we report on progress in folate biofortification and discuss the agroeconomical aspect of biofortified crop plants.

Adapted from:

Blancquaert D, Storozhenko S, Loizeau K, De Steur H, De Brouwer V, Viaene J, Ravanel S, Rébeillé F, Lambert W, Van Der Straeten D. 2010. Folates and folic acid: from fundamental research toward sustainable health. Critical Reviews in Plant Sciences **29**, 14-35.

Introduction

Tetrahydrofolate and derivatives thereof, collectively called folates, are tripartite water-soluble B-vitamins. The term 'folate' is derived from the Latin for leaf, 'folium', the leaf indeed being quite a rich source of folates. Folates are important cofactors in C1 metabolism, playing a role as one-carbon donors and acceptors (Scott *et al*., 2000; Rébeillé *et al*., 2006; Bekaert *et al*., 2008). Thus, they are involved in the formation of purines and thymidylate for DNA synthesis, but are also implicated in methionine, serine and glycine biosynthesis and in histidine catabolism (Gao *et al.*, 1998; Hennig *et al.*, 1998; Cook, 2001). Furthermore, they are important for the synthesis of pantothenate (vitamin B5). Finally, folates play a central role in the methyl cycle, which supplies a vast number of methylation reactions through S-adenosylmethionine. In plants these are of primary importance in lignin, alkaloid, betaine, and chlorophyll biosynthesis, besides being essential in photorespiration. Recently, a role of folate in plant N metabolism (Jiang *et al*., 2013), auxin signaling (Stokes *et al*., 2013) and chromatin silencing (Zhou *et al*., 2013) has been reported.

Folate deficiency in humans can result in the onset of several diseases and disorders. A causal link between low folate status and the occurrence of megaloblastic anemia and neural tube defects (NTD), such as spina bifida and anencephaly, was convincingly proven (Cherian *et al*., 2005; Li *et al*., 2006b; Berry and Li, 2002; Rader and Schneeman, 2006; De Wals *et al*., 2007). Low folate status is also associated with the occurrence of several neurodegenerative disorders, including Alzheimer's disease (Seshadri *et al.,* 2002), a higher risk of cardio-vascular disease (Stanger, 2004), the development of a range of cancers (Choi and Friso, 2005) and major depressive disorder (Papakostas *et al*., 2012); however, no causal relationship has been established for any of these so far.

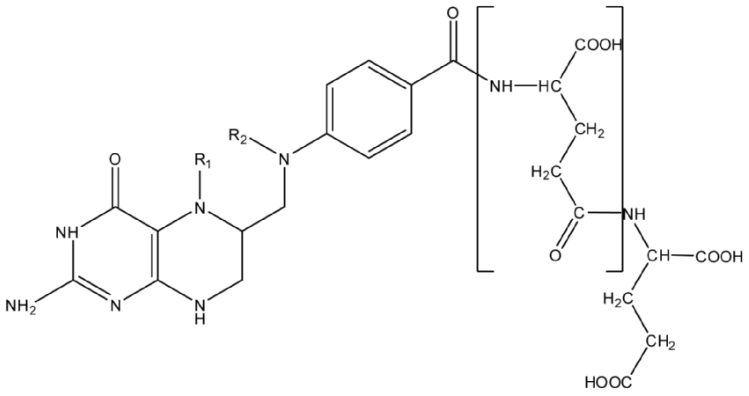
Humans cannot synthesize folates *de novo* and therefore depend entirely on their dietary sources. Most staple crops, such as potatoes, bananas and cereals are poor sources of folates (USDA National Nutrient Database for Standard References; http://www.nal.usda.gov/fnic/foodcomp/search/), which makes folate deficiency a global problem. Recently, folate biofortified tomatoes and rice have been developed, a major leap forward in fighting folate deficiency (Diaz de la Garza *et al*., 2007; Storozhenko *et al*., 2007b; Bekaert *et al*., 2008). However, folates are labile compounds, sensitive to spontaneous and photo-oxidative degradation. Consequently, after harvesting and during food storage and processing, folate losses can be massive. For example, in tomato fruits and *Arabidopsis* leaves, a daily loss of approximately 10% of total folate was detected (Orsomando *et al*., 2006), whereas the breakdown rates in humans reach only 0,5% per day (Gregory and Quinlivan, 2002). Folate salvage and the enhancement of folate stability can offer a solution and therefore recently gained interest.

Chemical properties of folates

Folate molecules consist of the following three moieties: a pteridine ring, a *para*-aminobenzoate (*p*-ABA) ring and a tail of one or more L-glutamates (Figure 1), linked through their γ-carboxyl groups. The folate pool of a cell or a food source is a mixture of related molecules, which differ in their oxidation state, in the attached one-carbon (C1) unit and in the length of the glutamate tail.

Tetrahydrofolate is the most reduced folate and is the only active cofactor, whereas folic acid (FA), a form that does not occur naturally, is fully oxidized and hence, the most stable folate form.

The C1 substituents attached to the N5 position of the pteridine and/or to the N10 position of *p*-ABA determine the identity, properties and functions of the different folates (Rébeillé *et al.*, 2006). The C1 units occur as formyl (O=CH-), methylene (-CH_2-), methenyl (-CH=), methyl (CH_3-) or formimino (HN=CH-) group, each playing a specific role in C1 metabolism. Substitution on N-5 gives greater stability. Depending on the source, the length of the glutamate tail varies between one to eight L-glutamate molecules (Ravanel *et al.*, 2011). In animals, monoglutamates can only be found in the following two places: in the blood plasma and near the brush border of the intestine, where folate polyglutamates are deconjugated to monoglutamates by a pteroylpolyglutamate hydrolase (usually simply referred to as 'conjugase') prior to absorption. The polyglutamate tail retains folates within cells or cellular compartments, but also determines protein binding and hence enzyme activity, which is proportionate with the chain length (Appling, 1991; Shane, 1989). The ratio between folate monoglutamates and polyglutamates varies in plants (McGuire and Coward, 1984; Imeson *et al.*, 1990) and is, for instance, approximately 50% in rice seeds (Storozhenko *et al.*, 2007b).



R1	R2	
H	H	THF
CH3	H	5 - methyl
CHO	H	5 - formyl
CH=NH	H	5 - formimino
H	CHO	10 - formyl
- CH2 -		5,10 - methylene
= CH -		5,10 - methenyl

Figure 1 Chemical structure of folates. Folates are tripartite molecules, consisting of a pteridine moiety, *p*-ABA and a glutamate tail (a glutamate unit is represented between brackets). They differ in the extent of polyglutamylation, oxidation state and type of C1 unit attached (R1, R2).

Folic acid is commercially produced and can be found in supplements (folate pills) and fortified food such as breakfast cereals and flour. When FA is taken up by the body in the form of folate pills, it is readily reduced to DHF and THF by dihydrofolate reductase in the gut mucosa, thus becoming biologically active. It is then further converted to 5-methyl-THF (Scott *et al.*, 2000). (see section: Folate biosynthesis).

Folate functions

Folate synthesis and interconversions, nucleic acid biosynthesis and the methyl cycle form a complex network of reactions (Figure 2), all of which are part of C1 metabolism. In this network, folates play an important role as C1 donor and acceptor. Altogether, they are responsible for the biosynthesis of purines, thymidylate, methionine, pantothenate and formyl methionyl tRNA, serine and glycine interconversion and histidine catabolism (Fowler, 2001).

Purines receive their carbon two and carbon eight from 10-formyltetrahydrofolate, which is converted to tetrahydrofolate. 10-Formyltetrahydrofolate is also involved in the synthesis of formyl-methionine-tRNA. Thymidylate synthase uses 5,10-methylenetetrahydrofolate as C1 donor to convert a uracil type base to a thymine base. This reaction results in the formation of dihydrofolate, which can be reduced back to tetrahydrofolate by dihydrofolate reductase (DHFR; see section Folate biosynthesis). The biosynthesis of pantothenate (vitamin B5), a precursor of co-enzyme A, requires 5,10-methylenetetrahydrofolate as well.

5,10-Methylenetetrahydrofolate can be reduced to form 5-methyltetrahydrofolate which enters the methyl cycle. This reduction is mediated by 5,10-methylenetetrahydrofolate reductase, an enzyme which was first purified from pig and human liver (Daubner *et al.*, 1982; Zhou *et al.*, 1990). In mammals, this reaction is NADPH-dependent, almost irreversible and feedback inhibited by S-adenosylmethionine (SAM), whereas in plants it is NADH-dependent, reversible and not regulated by SAM (Roje *et al*., 1999). 5-Methyl-THF donates a methyl group to homocysteine to produce methionine. This conversion is catalyzed by the cobalamin (vitamin B12)-dependent methionine synthase. Methionine can be further converted to SAM by SAM synthetase, a reaction that requires ATP (Scott *et al.*, 2000). The methyl cycle supplies SAM to methyltransferases, enzymes which transfer methyl groups to a broad range of substrates, e.g. proteins, lipids, DNA, and hormones. This cycle also plays an active role in the degradation of excess dietary methionine (Scott *et al.*, 2000). Methylation also affects gene expression through DNA (Razin and Riggs, 1980; Tate and Bird, 1993) and histone modification (Allfrey *et al*., 1964; Kouzarides, 2002).

Besides their above-mentioned role in methionine metabolism, folates are also required for the metabolism of other amino acids, such as glycine, serine, and histidine. Tetrahydrofolate acts as a C1 acceptor in the conversion of serine to glycine. This reaction results in the formation of 5, 10-methylenetetrahydrofolate, which is an important source of C1 units. The reaction is catalyzed by serine hydroxymethyltransferase (SHMT). On the other hand, in mitochondria, glycine oxidation to CO_2 and NH_3 by the glycine decarboxylation complex (GDC) is also a potential source of C1 units. In addition, this reaction is important for photorespiration in plants (see below).

The catabolism of histidine is a process of four reactions wherein the last step requires folates. After conversion to formimino glutamate, the formimino group is transferred to tetrahydrofolate, yielding 5-formimino-tetrahydrofolate. Subsequently, the latter is converted to 5,10-methenyltetrahydrofolate. These two reactions are mediated by formiminotransferase cyclodeaminase, an octameric bifunctional enzyme residing in the Golgi apparatus (Gao *et al.*, 1998; Hennig *et al.*, 1998). However, the existence of these reactions in plants remains to be proven.

In plants, folates are also necessary in the biosynthesis of chlorophylls and lignin, through methylation, and play an important role in photorespiration (Hanson and Roje, 2001; Douce *et al*., 2001). Glycine formed in the peroxisomes is converted to serine in mitochondria. This process relies on two folate-dependent enzymes, the glycine decarboxylase complex GDC and SHMT. THF is a

cofactor of the T-protein of GDC and is converted to 5,10-CH2-THF upon glycine oxidation by GDC. THF is recycled by the reverse reaction, forming serine from glycine, which is catalyzed by SHMT (Rébeillé *et al*, 1994; Douce *et al*, 2001). Both GDC and SHMT are light-induced and accumulate to up to 40% of soluble protein in the mitochondrial matrix (Douce *et al*, 2001). Recently it was shown that folates act together with sucrose to influence auxin sensitivity and auxin distribution during seedling development (Stokes *et al*., 2013).

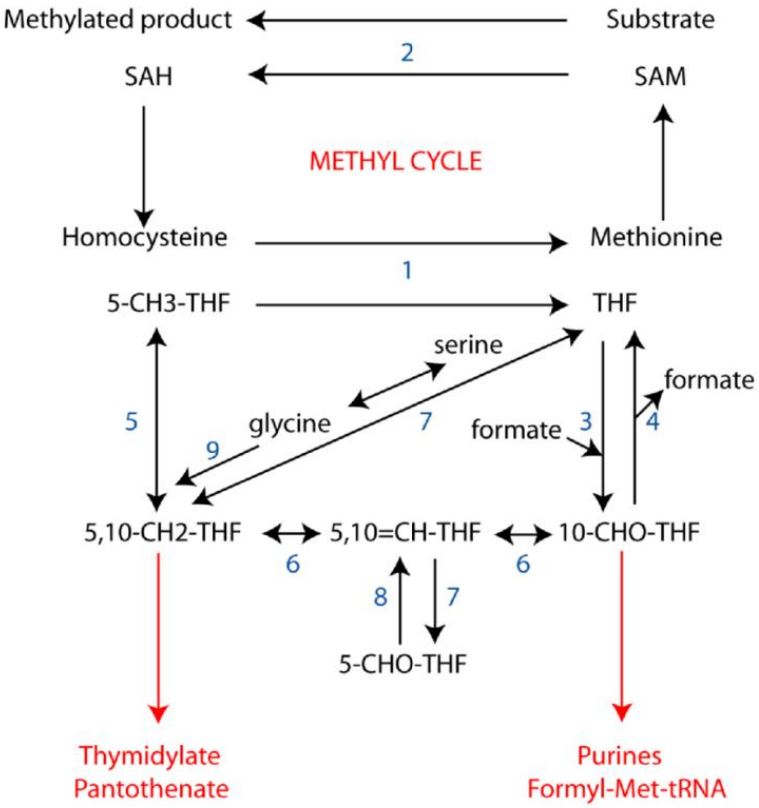


Figure 2. Schematic representation of folate metabolic pathways, interconversions and functions (in red). Abbreviations: SAH, S-adenosylhomocysteine; SAM, S-adenosylmethionine; THF, tetrahydrofolate; 5-CH_3-THF, 5-methyltetrahydrofolate; 5,10-CH_2-THF, 5,10-methylenetetrahydrofolate; 5,10=CH-THF, 5,10-methenyltetrahydrofolate; 10-CHO-THF, 10-formyltetrahydrofolate; 5-CHO-THF, 5-formyltetrahydrofolate. Enzymes: 1, methionine synthase; 2, methyltransferases; 3, 10-formyltetrahydrofolate synthase; 4, 10-formyltetrahydrofolate deformylase; 5, 5, 10-methylenetetrahydrofolate reductase; 6, 5,10-methylenetetrahydrofolate dehydrogenase/5,10-methenyltetrahydrofolate cyclohydrolase; 7, serine hydroxymethyl transferase; 8, 5-formyltetrahydrofolate cycloligase; 9, glycine decarboxylase complex.

Sources of food folate

Folate levels vary greatly between food sources, the highest being found in yeast, eggs, liver, green leafy vegetables, leguminous vegetables and certain fruits (e.g. oranges) (Table 1). Important staples, for instance rice, wheat, potatoes and bananas, have very low folate concentrations (USDA National Nutrient Database for Standard References; http://www.nal.usda.gov/fnic/foodcomp/search/). Fermented products such as beers, cheese and yoghurt are also important folate sources. However,

it is important to keep in mind that alcohol decreases intestinal folate absorption. In plant food, the predominant folate form found is 5-methyltetrahydrofolate (Holland *et al.*, 1996).

Methods of harvest and post-harvest treatment and storage of plant foods determine the consumable folate concentration. Folate biosynthesis continues after harvesting, but may decline rapidly (Orsomando *et al*., 2006). Moreover, physical damage and exposure to environmental factors, such as heat, oxygen and pressure during and after harvesting further decrease folate concentration present in plant foods. Ideally, folate sources need to be stored at low (freezing) temperature, with a low oxygen tension and in a dark environment to maintain maximal folate levels (Scott *et al.*, 2000). In addition, a short interval to market, sale and consumption are necessary to protect the folate pool present in its sources.

The effects of processing have been extensively reviewed (Scott *et al.*, 2000). Since folates are water soluble B-vitamins, great losses occur during cooking, blanching, washing and canning. Experiments in biofortified rice showed a folate loss of 45% after cooking for thirty minutes (Storozhenko *et al.*, 2007b). In pea, microwave cooking did not result in a folate loss, whereas 25 % were lost after boiling for ten minutes (Pitcher and Edwards, 1995).

Source	Folate content (µg/ 100g FW)
Soybean	370
Spinach	190
Chick pea	180
Broccoli	90
Brown rice	20
White rice, cooked	3
GA rice, uncooked	1700
GA rice, cooked	850

Table 1. A few examples of folate rich vegetables. As a comparison, rice (wt rice and transgenic GA rice) is added. Approximately 50% of folates are lost after cooking GA rice (Storozhenko *et al*., 2007). Values (except those for GA rice) were obtained from the USDA National Nutrient Database for Standard Reference, Release 25 (USDA, 2012).

Folate bioavailability

Folate bioavailability is defined as the portion of folates present in a diet, which eventually can be used in folate-requiring processes. This has been thoroughly reviewed by Gregory and Scott (Gregory, 1995; Gregory, 1997; Scott *et al.*, 2000). Folate bioavailability is the ultimate summation of the following: (1) release of folates from the food matrix, (2) uptake by the brush border, (3) deconjugation (mediated by conjugase, which turns polyglutamates into monoglutamates), (4) active transport and diffusion, (5) conversion to 5-methyltetrahydrofolate, (6) release into the circulation system and (7) transport into the different tissues (figure 3) (Scott *et al*., 2000). Bioavailability also varies between folates, e.g. folic acid is 70% more bioavailable than natural occurring folates (Food and Drug Administration, 1996; Bailey, 2004). Although folate bioavailability is of paramount

importance to measure the effect of fortification and biofortification, its study is extremely complex, considering the different factors that play a role in its determination.

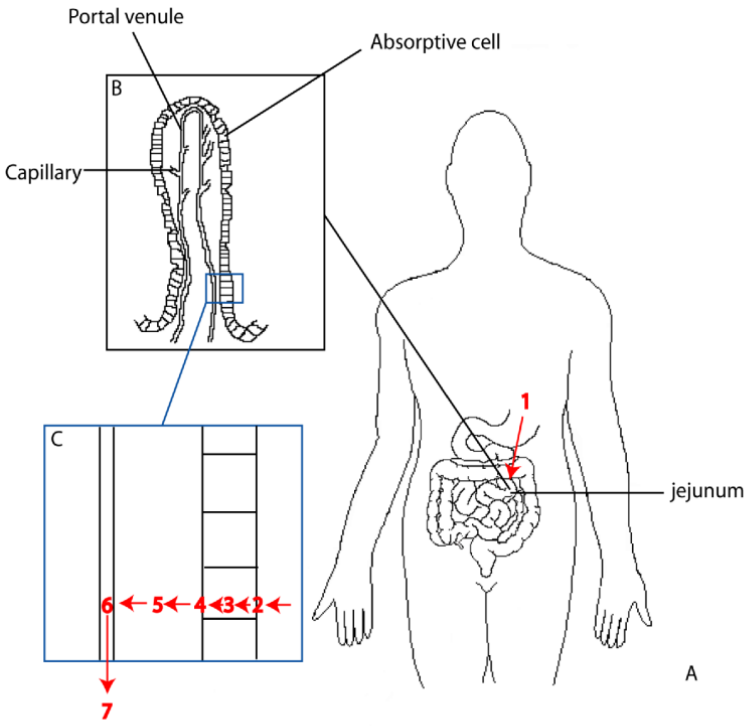


Figure 3. The different aspects of folate bioavailability. Seven steps are crucial in its determination (1 to 7 on figures A and C): (1) release from the food matrix, (2) uptake by the jejunal brush border, (3) conversion of polyglutamates to monoglutamates in the brush border membrane by deconjugase, (4) transport and diffusion into the inner layers of the jejunal mucosa, (5) conversion to 5-methyltetrahydrofolate (in the mucosal enterocytes), (6) release into the circulation system and (7) transport into the different tissues. (A) schematic representation of the human body and its digestive system, (B) cross-section of a jejunal villus, (C) enlargement of the mucosa and brush border.

Folate deficiency

Since folates are involved in nucleic acid biosynthesis and the methyl cycle, folate deficiency results in malfunctioning of these vital processes. Hence, inadequate folate intake may lead to the onset of diseases and disorders, such as megaloblastic anaemia (a form of macrocytic anemia caused by folate and/or vitamin B12 deficiency), neural tube defects (NTDs) (Geisel, 2003) and a series of neurodegenerative disorders, amongst which Alzheimer's disease and other forms of dementia are the best known examples (Seshadri *et al.*, 2002). Folate deficiency was also correlated with a higher risk of coronary and cardiovascular diseases (Scott and Weir, 1996; Quinlivan *et al.*, 2002; Stanger, 2004), stroke (Wald *et al.*, 1998), major depressive disorder (Papakostas *et al.*, 2012) and a wide range of cancers, such as leukemia, colorectal, breast, cervical, pancreatic and bronchial cancer (Lucock *et al.*, 2003; Choi and Friso, 2005). It is important to mention that a causal relationship between folate deficiency and the occurrence of the disorder has only been convincingly shown for megaloblastic anemia and for NTDs such as spina bifida or anencephaly. Adequate dietary folate intake can prevent these conditions (Berry and Li, 2002; Rader and Schneeman, 2006; De Wals *et al.*, 2007).

Although the underlying mechanisms are in most cases not completely elucidated, these disorders can be divided by cause in three groups: homocysteinemia, hypomethylation and an errant DNA biosynthesis. Homocysteinemia, the presence of high homocysteine levels in the circulation system, is caused by an inadequate conversion of homocysteine to methionine (See Folate functions). Three

different enzymes control homocysteine levels in cells and blood plasma, two of which require folate as a cofactor, i.e. methionine synthase and methyleneTHF reductase (MTHFR) (Scholl and Johnson, 2000). Methylation is important in a wide range of reactions, from the stabilization of myelin basic protein, which plays a role in the synthesis of the lipid layer lining nerve axons, to hormone inactivation and gene expression control. Therefore, hypomethylation, caused by inadequate folate levels, can lead to changes in expression of oncogenes and/or neuropathy. Impaired DNA biosynthesis is most pronounced in rapidly dividing tissues, such as bone marrow, which require high folate concentrations. Thymidylate synthase (TS) is responsible for the conversion of a uracil base to a thymidine base. Low availability of folates may lead to a shortage of dTMP. Subsequently, dUTP may be incorporated in DNA, which becomes unstable. Ultimately, this instability may cause point mutations, single and double strand breaks and even chromosomal breakage.

Folate deficiency is a global problem, which is most pronounced in the less developed parts of the world. However, even in the developed world folate deficiency is a widespread phenomenon, leading to region-specific increased prevalence of related congenital disorders (EUROCAT Database http://www.bio-medical.co.uk/eurocatlive). The prevalence of NTDs, such as spina bifida and anencephaly, in Shanxi province, one of the poorest regions of China, is more than ten times higher (up to 199.38 in 10000 live births; Li *et al.*, 2006a; Gu *et al*., 2007) than in the Western World. A study in Balrampur District, the least-developed region in India, showed similar results (Cherian *et al.*, 2005). The recommended/reference daily intake (RDI) of folates is 400 µg for an adult. For pregnant women, this figure needs to be increased to 600 µg. However, studies indicate that most diets provide daily approximately 200 µg of folates (Scott *et al*., 2000; Ward *et al*., 1997).
Several strategies have been proposed to fight folate deficiency. Folic acid supplementation (capsule intake) has shown limited success in practice. The neural tube is formed 21-27 days after conception, before most women are aware of their pregnancy. Thus, in order to prevent NTDs, women should take folic acid supplements on a regular basis from the peri-conceptional phase until approximately 3 months of gestation. Moreover, a lot of pregnancies are unplanned. For instance, in the United States, this is the case for about half the pregnancies (Geisel, 2003), a number which is reduced to 44% in Europe (Singh *et al*., 2010). Another strategy to fight folate deficiency is the fortification of foods. Since 1998, it is compulsory in the United States and Canada to fortify cereal-based foods, such as flour and breakfast cereals, with folic acid (Food and Drug Administration, 1996). This resulted in a 20 to 53% drop of spina bifida incidents and up to 38% reduction in prevalence of anencephaly (Green, 2002; De Wals *et al*., 2007). Plasma homocysteine levels decreased after folate fortification (Jacques *et al*., 1999; Ray *et al*., 2000), but a Canadian study proved that this could be a temporary effect (Garcia *et al*., 2008). The amount of added folic acid is such that the predicted average intake resulting from consumption of fortified food products equals ~100 µg /day (corresponding to 170 µg Dietary Folate Equivalents (DFE)), since folic acid is assumed to be 1.7 times more bioavailable than natural folates (Food & Drug Administration, 1996; Bailey, 2004). Evaluation of efficacy showed that reality surpasses this prediction, and that the RDI was met or exceeded in most adults (Choumenkovitch *et al*., 2002; Quinlivan and Gregory, 2003). A recent study on the effect of the Canadian folic acid fortification program revealed that whole blood cell folate levels increased continuously after folate fortification was initiated and did not stabilize, not even after six years (Garcia *et al*., 2008). Very high folate intake, however, can also have certain adverse effects. Excessive intakes of folic acid (>1 mg/day) may mask the diagnosis of vitamin B12 deficiency (Mills *et al*., 2003; Ray *et al*., 2003). High folate status may also reduce the response to antifolate drugs used

against rheumatoid arthritis, psoriasis, and cancer (Smith *et al*., 2008) and results in cognitive decline in the elderly, especially in combination with low cobalamin levels (Morris *et al*., 2007). Moreover, folic acid supplementation to cobalamin deficient women may also impair fetal growth and brain development (Takimoto *et al*., 2011; Marean *et al*., 2011) and could be harmful to the nervous system when combined with vitamin B12 deficiency (Reynolds, 2006).Folic acid fortification does not protect against ischeamic heart disease (Wald *et al*., 2006) and has no effect on breast cancer risk (Lewis *et al*., 2006). The risk for colorectal cancer has been proven to increase with high plasma concentrations of folate (Lewis *et al*., 2006). For the prevention of cancers, it seems that the timing of folic acid supplementation is crucial. Trials on colorectal cancer in mice (Song *et al*., 2000a; Song *et al*., 2000b) and humans (Cole *et al*., 2005; Bonaa *et al*., 2006) support this statement. Therefore, fortification remains a controversial issue in the EU (Finglas *et al*., 2006; Hubner *et al*., 2007) even though incidence of NTDs varies by a factor of 6 over different regions in Europe (Eurocat Database). Moreover, doses of folic acid above 300 µg cannot be converted to 5-methyltetrahydrofolate in the mucosa, resulting in the presence of unaltered folic acid in the circulation system (Scott *et al.*, 2000). This may mask pernicious anaemia in vitamin B12 deficient cells and possibly promotes tumour growth (Scott *et al*., 2000). Food fortification also requires quality control and specialized infrastructure, which usually are very restricted in developing countries. Altogether, folic acid supplements and fortified foods are not sufficient to prevent folate deficiency, especially since in most cases rural populations cannot benefit thereof. A third solution to limit folate deficiency is the intake of folate-rich food. However, dietary changes have been proven to be insufficient to reach a daily folate intake of 600 µg (Scholl and Johnson, 2000).

Folate biofortification, the enhancement of folates in staple food through plant breeding or biotechnology, offers an alternative and complementary approach to the above-mentioned strategies. Since folic acid and natural folates have different effects on folate metabolism (Ross *et al*., 1984), the above-mentioned possible negative consequences of folic acid fortification may not be applicable on folate biofortification. Colorectal cancer risk, for instance, is proved to decrease with enhanced intake of dietary folate (unlike folic acid) (Sanjoaquin *et al*., 2005). In addition, 5-methyl-THF is likely not masking vitamin B12 deficiency in elderly, as opposed to folic acid (Gutstein *et al.,* 1973). Therefore, transgenic plant lines should be created with appropriate folate levels. Moreover, development of biofortified foods represents a single investment cost and seeds can be re-sown from last season's harvest. In addition, biofortified crops can reach the poorest populations after introduction of the traits into local varieties.

Folate biosynthesis

The *de novo* biosynthesis of folates is restricted to plants and micro-organisms. In plants, it is characterized by a compartmentation of the reactions in three subcellular locations (Figure 4). Pterins and *p*-ABA are first produced separately in, respectively, the cytosol and plastids and subsequently condensed, glutamylated and reduced to form tetrahydrofolate monoglutamate in the mitochondria (Ravanel *et al*., 2011). The monoglutamates become polyglutamylated by the action of mitochondrial folylpolyglutamyl synthase (FPGS). While polyglutamylation favors mitochondrial retention, folates are exported to the cytosol as monoglutamates. Plastids and cytosol also possess FPGS activities (Ravanel *et al*., 2001).

At present, all the enzymes participating in the biosynthesis of tetrahydrofolate in plants and bacteria have been –at least partially- characterized. A total of eleven steps are required to synthesize tetrahydrofolate, of which one step is mediated by a non-specific enzyme (see below). The reactions themselves are basically the same in plants as in bacteria and fungi (Hanson and Gregory, 2002).

Synthesis of 6-hydroxymethyldihydropterin

Pterins play an important role in pigments and the metabolism of some aromatic acids (Kaufman and Kaufman, 1985; Kuhn and Lovenberg, 1985; Johnson and Rajagopalan, 1985). In folate biosynthesis, the pteridine moiety, 6-hydroxymethyldihydropterin (HMDHP), is produced from GTP in four steps. None of the corresponding enzymes contain predicted targeting peptides, suggesting that they are cytosolic.

The first step comprises the conversion of GTP to dihydroneopterin triphosphate. This reaction is catalyzed by GTP cyclohydrolase I (GTPCHI), a dimer, of which each monomer consists of two GTPCHI domains in tandem. Both are essential to perform the activity (Basset *et al*., 2002). The primary structure of the plant protein is unique (Basset *et al.*, 2002). This enzyme is also present in mammalian cells, where it plays a role in the synthesis of tetrahydrobiopterin (Fan and Brown, 1976). A feedback inhibition on the plant enzyme has been suggested (Sohta *et al*., 1997); its mammalian counterpart is proven to be feedback sensitive (Yoneyama and Hatakeyama, 1998).

The second and third steps consist of the removal of the triphosphate chain to produce dihydroneopterin (DHN). This is done by two enzymatic actions. A nudix hydrolase (dihydroneopterin triphosphate pyrophosphatase) releases a pyrophosphate group (Klaus *et al.*, 2005a), while a nonspecific phosphatase cleaves the remaining phosphate from dihydroneopterin phosphate to yield DHN (Suzuki and Brown, 1974).

The fourth step towards HMDHP synthesis is mediated by DHN aldolase (DHNA), which cleaves the lateral side chain of DHN to release glycolaldehyde and HMDHP (Goyer *et al.*, 2004). The structure of this enzyme differs greatly between organisms. In plants, three isoforms exist (Goyer *et al*., 2004), whereas in fungi this activity is part of a trifunctional enzyme (together with HPPK and DHPS respectively, see below) (Guldener *et al.*, 2004). The bacterial protein appears to be an octamer (Mathis and Brown, 1980).

Synthesis of para-aminobenzoic acid

Besides folate biosynthesis and esterification to glucose, no other metabolic fate has been assigned to *p*-ABA in plants. *p*-ABA is formed in plastids from chorismate (Ravanel *et al*., 2011). This molecule originates from the shikimate pathway (Weaver and Herrmann, 1997) and is also produced in plastids (Hrazdina and Jensen, 1992). Like pterines, it plays an important role in the synthesis of aromatic amino acids (Siehl, 1998). The synthesis of *p*-ABA requires two steps (Nichols *et al*., 1989). First, 4-aminodeoxychorismate (ADC) and glutamate are produced from chorismate and glutamine by ADC synthase (Ye *et al*., 1990). This step comprises the following two reactions: the release of NH_3 from glutamine (glutaminase activity) and the substitution of NH_3 for the hydroxygroup at position 4 of chorismate (ADC synthase activity), during which NH_3 is most probably channeled throughout the activity domains in plants (Sahr *et al.*, 2006). The amination activity is inhibited by ADC (Camara *et al*., 2011). Subsequently, pyruvate is removed and aromatization takes place to form *p*-ABA. Although it can be spontaneously, but slowly, aromatized to *p*-ABA (Tewari *et al.*, 2002), this esterification is mediated by ADC lyase (Ye *et al*., 1990).

In bacteria, three genes encode PabA, PabB and PabC proteins. The first two assemble to form a heterodimer with ADC synthase activity. The third contains pyridoxal phosphate and displays ADC lyase action (Green *et al.*, 1992). Hybrid proteins, consisting of fused domains homologous to bacterial PabA and PabB proteins, were found in plants, fungi and *Plasmodium*. These proteins form ADC and should be considered ADC synthases (Basset *et al.*, 2004a). In plants, the ADC synthase protein exists as a bipartite fusion protein, with tandem domains homologous to the bacterial subunits (Basset *et al.*, 2004a). This monomeric enzyme (95 kDa) is inhibited by dihydrofolate and methotrexate (unlike prokaryotic ADC synthases), suggesting it could be a target for antifolate drugs (Sahr *et al.*, 2006). A homodimeric ADC lyase has also been characterized in plants (Basset *et al.*, 2004b). ADC synthase and ADC lyase are located in chloroplasts and show no feedback inhibition by *p*-ABA, nor by its glucose ester or by folates (tetrahydrofolate, 5-methyltetrahydrofolate and 5-formyltetrahydrofolate pentaglutamate) (Basset *et al.*, 2004a; Basset *et al.*, 2004b).

p-ABA is a hydrophobic, weak acid, which is membrane-permeable in its uncharged form and is subject to a pH-dependent subcellular distribution as an anion. This feature is not desirable, most importantly because it acts as an inhibitor of dihydropteroate synthase activity (DHPS, see below). However, in plants, *p*-ABA is reversibly converted to its β-D-glucopyranosylester by a UDP-glucose:*p*-ABA glucosyltransferase (*p*AGT) activity (Quinlivan *et al.*, 2003). This activity is most probably mediated by a single isoform (Eudes *et al.*, 2008). Since most of free *p*-ABA is esterified in the cell and approximately 88% of this glucose ester is present in vacuoles, it was suggested to act as a *p*-ABA storage form (Quinlivan *et al.*, 2003; Eudes *et al.*, 2008).

Synthesis of dihydrofolate

Biosynthesis of dihydrofolate in plants occurs solely in mitochondria (Neuburger *et al.*, 1996; Ravanel *et al.*, 2001). Three reactions are necessary to convert HMDHP and *p*-ABA into dihydrofolate. First, HMDHP is activated as a pyrophosphate compound by dihydropterin pyrophosphokinase (HPPK). This activation requires ATP. Subsequently, HMDHP-PPi is condensed with *p*-ABA to produce dihydropteroate. This condensation is catalyzed by dihydropteroate synthase (DHPS). In *E. coli*, HPPK and DHPS activities are performed by two monofunctional proteins (Dallas *et al.*, 1992); whereas in eukaryotes, different situations are present. *Plasmodium* contains a bifunctional HPPK/DHPS enzyme (Triglia and Cowman, 1994), fungi have trifunctional HPPK/DHPS/DHNA proteins (Guldener *et al.*, 2004) and plants possess polypeptides with fused HPPK/DHPS domains and a mitochondrial transit peptide (Rébeillé *et al.*, 1997). In plants, similar rates of AMP and DHP synthesis were found in the presence of *p*-ABA indicating a good coupling between the two reactions. Plant DHPS domains are a putative regulatory point of folate biosynthesis, since they are feedback inhibited by dihydropteroate, dihydrofolate monoglutamate and tetrahydrofolate monoglutamate (Mouillon *et al.*, 2002). In *Arabidopsis*, a second, cytosolic HPPK/DHPS isoform was found, highly abundant in developing seeds, where it may play a folate independent role in stress response (Storozhenko *et al.*, 2007a; Navarrete *et al*., 2012).

The last step to synthesize dihydrofolate is the attachment of a first glutamate to the carboxy part of the *p*-ABA moiety of dihydropteroate. This ATP-dependent reaction is mediated by dihydrofolate synthetase (DHFS). Tetrahydrofolate polyglutamate synthesizing bacteria have bifunctional DHFS/FPGS enzymes (Bognar *et al.*, 1985), whereas in fungi and higher plants a monofunctional monomeric DHFS is present, which is encoded by a single gene in *Arabidopsis* (Ravanel *et al.*, 2001). Animals lack the three aforementioned enzymes, together with ADC synthase, ADC lyase and DHNA, making them incapable of synthesizing folates *de novo*.

Synthesis of tetrahydrofolate

From dihydrofolate, two steps remain towards polyglutamylated folates. First, dihydrofolate is reduced to tetrahydrofolate by dihydrofolate reductase (DHFR), which uses NADPH as an electron donor. Second, glutamate residues are attached to the glutamate moiety in THF in an ATP-dependent reaction, catalyzed by folylpolyglutamate synthetase (FPGS).

A monomeric DHFR protein is present in bacteria, yeast and vertebrates, while algae and protozoa possess a bifunctional homodimeric DHFR/TS (thymidylate synthase) enzyme. In plants, DHFR was described either as a bifunctional homodimeric enzyme (Cella and Parisi, 1993) or as a monofunctional polypeptide associated with TS as part of a multimeric complex (Toth *et al*., 1987). TS catalyses the conversion of deoxyuridine-monophosphate (dUMP) and 5,10-methylenetetrahydrofolate into deoxythymidine-monophosphate (dTMP) and dihydrofolate. Subsequently, dihydrofolate is reduced to tetrahydrofolate by the DHFR activity of the enzyme. Plastidial and mitochondrial DHFR isoforms were respectively found in carrot and pea (Luo *et al.*, 1997; Neuburger *et al.*, 1996). In *Arabidopsis*, 3 isoforms exist, one of which is predicted to be cytosolic (Storozhenko *et al*., 2005). Likewise, a cytosolic isoform may occur in rice. DHFR is a target for anti-folate (anti-cancer) drugs such as methotrexate (Scott *et al*., 2000).

Bacterial FPGS proteins, encoded by the *FolC* gene, also exhibit DHFS activity. However, in eukaryotes, the enzyme is monofunctional. The dicot model plant *Arabidopsis* has three isoforms of the protein: a plastidial, cytosolic and mitochondrial FPGS, each encoded by a unique gene. They display broad substrate specificity (adding glutamates to H4F-Glu_1 carrying various C1 units; Ravanel *et al.*, 2001). Reports on fpgs1 mutants, encoding the plastidial isoform, suggested a role of this enzyme in post-embryonic root development (Srivastava *et al*., 2011) and global DNA methylation and histone H3K9 dimethylation (Zhou *et al*., 2013) in *Arabidopsis*. In mammals and fungi, only one gene is present, encoding, depending on the translation initiation codon, cytosolic and mitochondrial isoforms (Freemantle *et al.*, 1995; DeSouza *et al.*, 2000).

Glutamylation is important to retain folates within cells or cellular compartments, since this reaction increases the anionic nature of folates (Appling, 1991). Polyglutamylated folates are also preferred above monoglutamate forms by folate-dependent enzymes (Shane, 1989), with the exception of folate transporters. Since plant cells can take up externally supplied folates (Crosti *et al.*, 1993) and folates occur in most plant cell compartments, carrier-mediated folate transport is suggested to be of great importance. To date, only one vacuolar and two plastidial transporters have been identified (Bedhomme *et al.*, 2005; Klaus *et al.*, 2005b; Raichaudhuri *et al*., 2009), whereas mitochondrial transporters are believed to exist as well. The vacuolar carrier, AtMRP1, is a multidrug resistance associated protein (MRP), belonging to the ATP-binding cassette (ABC) family of transporters (Raichaudhuri *et al.,* 2009). Although it was first suggested that a vacuolar carrier might transport polyglutamate forms (Orsomando *et al*., 2005), the uptake of folate polyglutamates by AtMRP1 and its equivalents is negligable (Raichaudhuri *et al.,* 2009). Since polyglutamylated folates are largely present in this compartment which lacks FPGS activity, the possibility of a non-MRP-type polyglutamylate folate transporter needs to be considered. In this case, plant vacuoles would import polyglutamylated folates, hydrolyze them, and export them as monoglutamylates, as demonstrated for mammalian lysosomes.

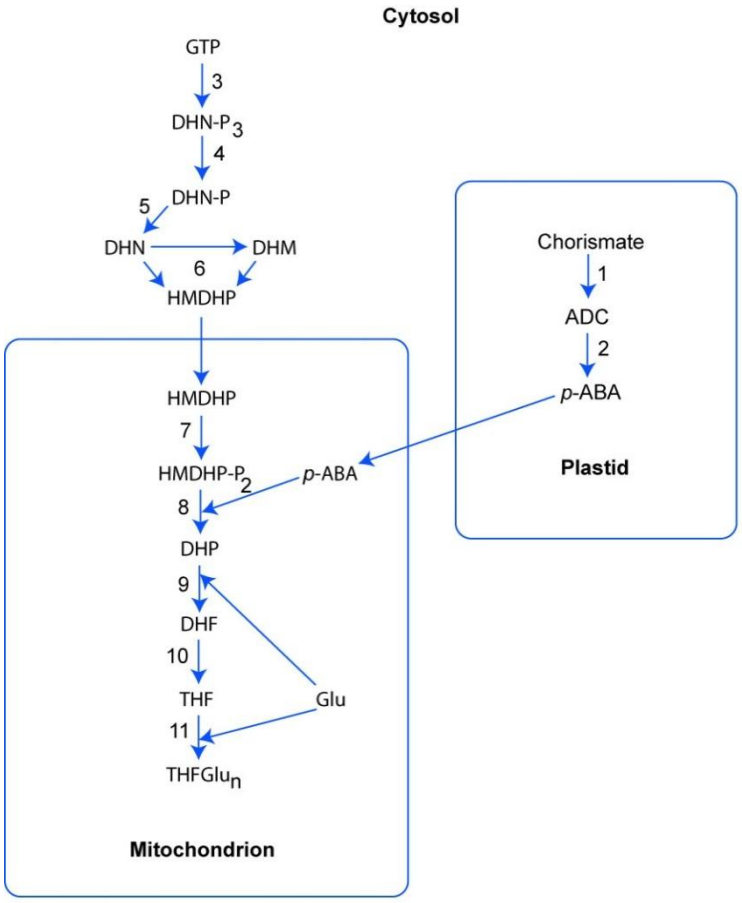


Figure 4. Schematic representation of folate biosynthesis, characterized by its compartmentalization. The pteridine moiety and *p*-ABA are synthesized in the cytosol and in the chloroplast respectively. After import into the mitochondria, they are condensed, glutamylated and reduced to form tetrahydrofolate monoglutamate. Polyglutamylation takes place in the cytosol, chloroplast and mitochondrion. Abbreviations: GTP, guanosine triphosphate; DHN-P_3, dihydroneopterin triphosphate; DHN-P, dihydroneopterin monophosphate; DHN, dihydroneopterin; DHM, dihydromonapterin; HMDHP, 6-hydroxymethyldihydropterin; HMDHP-P_2, 6-hydroxymethyldihydropterin pyrophosphate; DHP, dihydropteroate; DHF, dihydrofolate; THF, tetrahydrofolate; $THFGlu_n$, tetrahydrofolate polyglutamate; Glu, glutamate; ADC, aminodeoxychorismate; *p*-ABA, *para*-aminobenzoate. Enzymes: 1, ADC synthase; 2, ADC lyase; 3, GTP cyclohydrolase 1; 4, dihydroneopterin triphosphate pyrophosphatase; 5, non-specific phosphatase; 6, DHN aldolase; 7, dihydropterin pyrophosphokinase; 8, dihydropteroate synthase; 9, dihydrofolate synthetase; 10, dihydrofolate reductase; 11, folylpolyglutamate synthetase.

Folate distribution and plant development

Folate concentrations not only greatly differ amongst plant species, but also between plant tissues. Leafy vegetables contain more folates (1.5-4.5 nmol/g fresh weight) than rooty vegetables (0.3 nmol/g fresh weight) and fruits (0.2-0.8 nmol/g fresh weight) (Scott *et al*., 2000). The folate pool also decreases during fruit ripening and post-harvest (Scott *et al*., 2000; Basset *et al*., 2004a). In seeds of pea, 5-formyltetrahydrofolate is the most dominant folate derivative, while in other tissues 5-methyltetrahydrofolate reaches the highest concentration (Shin *et al*., 1975; Cossins, 2000). In rice

seeds however, 5-methyltetrahydrofolate accumulated most, accounting for up to 89% of the total folate pool (Storozhenko *et al*., 2007b).

Folates are present in all cell organelles, but are not equally distributed. Fractionation experiments have revealed the intracellular distribution of folates, with approximately 30% of total folate pool present in mitochondria, 40% in the cytosol, approximately 20% in vacuoles and only 10% in plastids (Jabrin *et al*., 2003; Orsomando *et al*., 2005). This distribution pattern becomes more complex when considering the different folate derivatives (Chen *et al*, 1997). In higher plants, 45-65% of the total folate pool consists of methyl derivatives, 10-15% of tetrahydrofolates and methylene derivatives, and 30-55% of formyl derivatives (Cossins, 2000). The mitochondrial pool is dominated by 5-formyl-THF, which is not a C1 donor but may act as an inhibitor of folate-dependent enzymes (Orsomando *et al*, 2005; Goyer *et al*, 2005). The chloroplastic pool is rich in 10-formyl and 5-methyl derivatives (Orsomando *et al*, 2005), in agreement with the involvement of this compartment in the *de novo* synthesis of purines and methionine, respectively (Zrenner *et al*, 2006; Ravanel *et al*, 2004). Also, methyl derivatives are predominant in the cytosol (Chen *et al*, 1997), where they are involved in the synthesis of methionine (Ravanel *et al*, 2004), and in the vacuole where they are possible candidates for folate storage (Orsomando *et al*, 2005).
Most of the folates present *in vivo* are polyglutamylated, but the values differ depending on tissue and organelle type. For instance, in the matrix of pea leaf mitochondria 25% of the polyglutamates are tetraglutamates and 55% pentaglutamates (Besson *et al*., 1993).
The actual amount of folates logically corresponds to the needs to supply C1 metabolism. Therefore, high folate concentrations are found in actively dividing tissues. For example, meristematic tissues of the root tip contain 5-fold more folate than the mature root (Jabrin *et al*, 2003) and *Arabidopsis* cell-suspension cultures, which have short generation time, contain up to 18 nmol/g fresh weight of folate (Loizeau *et al.*, 2007 and 2008). The pool of folate is also markedly increased during germination (Roos and Cossins, 1971), consistent with the transition of seeds from a quiescent to an active stage and the resumption of cell cycle and metabolic activities (Bewley, 1997). In germinating pea seeds, the changes in folate pool are more important in the embryo than in cotyledons, where the original amount of folate, which is already high (1 nmol/g fresh weight), triples during embryonic development. This trend is concordant with the accumulation of DHFR-TS and HPPK-DHPS mRNA's and proteins (Jabrin *et al*., 2003), indicating that this initial rise corresponds to a *de novo* synthesis of the cofactor. In the presence of sulfanilamide, an inhibitor of THF biosynthesis, sufficient folate is present in the seed to achieve germination *sensu stricto* but seedling growth is further blocked, indicating therefore that *de novo* synthesis of folates is crucial during the early stages of plant development (Gambonnet *et al*., 2001).

Light is another important factor that influences folate accumulation and distribution within plant cells. Indeed, green mature leaves are characterized by a high folate content that is triggered by the acquisition of photosynthesis and thus is related to light (Jabrin *et al*, 2003). The relationship between folate accumulation in leaves and light is not yet fully understood but the photorespiratory and methylation cycles have been shown to be responsible, at least in part, to the increased folate pool in green versus etiolated leaves. First, photorespiration involves two folate-dependent enzymes, GDC and SHMT, which accumulate within the mitochondria during greening (Douce *et al*, 2001). Part of the folate synthesized in light might contribute to maintain a high photorespiratory activity. This statement is supported by the similar expression patterns observed for genes coding for folate-

synthesizing enzymes (HPPK-DHPS and DHFR-TS) or enzymes involved in photosynthesis (small subunit of Rubisco) and photorespiration (the T, P and H proteins of GDC) during leaf development in the light (Vauclare *et al*, 1996; Jabrin *et al*, 2003). The subcellular distribution of folates is also affected by light, with a more important accumulation of the cofactor in the cytosol (5-fold) than in mitochondria (two-fold) during greening (Gambonnet *et al*, 2001; Jabrin *et al*, 2003). As mentioned above, the cytosolic pool of folates is dominated by 5-methyl-THF, thus suggesting an important turnover of SAM in photosynthetic leaves. The importance of methylation reactions in these tissues is well exemplified by chlorophylls that accumulate in thylakoid membranes and require a methylation step for their synthesis (Block *et al*, 2002). Thus, when etiolated leaves are treated with the antifolate methotrexate to reduce the folate pool, the rate of chlorophyll synthesis during illumination is reduced by 2.5-fold (Van Wilder *et al*, 2009). In these conditions, the modest reduction (25%) of the folate pool is accompanied by an important impairment of the methylation cycle, which is then responsible for a 3-fold reduction of the methyltransferase activity involved in chlorophyll synthesis (Van Wilder *et al*, 2009).

Together, the above presented data illustrate the crucial roles of folates in many aspects of plant biochemistry, physiology and development. These elements must be taken into consideration when developing a strategy to modify folate status in plants. Indeed, even moderate shortage in folate can dramatically affect plant development (Gambonnet *et al*, 2001; Van Wilder *et al*, 2009) and thus, directly have impact on plant health and productivity. Several recent studies indicated that the equilibrium between the different THF derivatives rather than the overall folate pool is important to allow a normal activity of C1 metabolism. For example, the short term adaptive response of *Arabidopsis* cells to folate depletion is characterized by a reorientation of C1 units towards the synthesis of nucleotides at the expense of methylation reactions (Loizeau *et al*, 2008). This rapid adaptation of folate homeostasis is controlled at the transcriptional level but in case of prolonged folate deficiency a posttranslational regulatory process can restore the flux of C1 units towards the synthesis of methionine and S-adenosylmethionine (Loizeau *et al*, 2007). Although counterintuitive, an overall folate increase can also affect plant metabolism and development because of important changes in folates distribution. Thus, *Arabidopsis* plants knocked out in the gene coding 5-formyl-THF cycloligase are characterized by an overall two-fold increase in folate pool, with an eight-fold increase in mitochondrial 5-formyl-THF and an extramitochondrial four-fold increase in 10-formyl/5,10-methenyl-THF (Goyer *et al*, 2005). As a consequence of 5-formyl-THF accumulation, the mitochondrial SHMT activity that is involved in photorespiration was inhibited, leading to a very large accumulation (46-fold) of glycine in leaves. This marked impairment of photorespiration was associated with a reduced growth rate and a delayed flowering of the mutant as compared to wild-type plants (Goyer *et al*, 2005). In addition, accumulation of 5-formyltetrahydrofolate, might inhibit the activity of numerous enzymes (Stover and Schirch, 1993; Roje *et al*., 2002). To conclude, plants can adapt to perturbation of folate homeostasis through transcriptional and/or post-transcriptional processes but these adjustments of C1 metabolism can have detrimental effect on plant growth and reproduction.

Folate stability

Folates are susceptible to spontaneous and photo-oxidative degradation by cleavage of the C9-N10 bond, bearing, depending on the source, tetra/dihydropterin-6-aldehyde and *para*-aminobenzoylglutamate (Suh *et al.*, 2001). Ring oxidation of tetrahydropterin-6-aldehyde (THPA) produces dihydropterin-6-aldehyde (DHPA) and subsequently the fully oxidized, aromatic form pterin-6-aldehyde (PA) (Hillcoat *et al.*, 1967; Whiteley *et al.*, 1968; Reed and Archer, 1980; Gregory, 1989), which can be further oxidized to pterin-6-carboxylate (Lowry *et al.*, 1949). 10-Formyltetrahydrofolate oxidizes to 10-formyldihydrofolate. Further oxidation leads to the biologically inactive form 10-formyl folic acid (Blackley, 1969).
Folates vary in their sensitivity to breakdown: tetrahydrofolate and dihydrofolate being the most cleavage-prone (Hillcoat *et al.*, 1967; Reed and Archer, 1980; Gregory, 1989), folic acid and 5-formyltetrahydrofolate being the most stable compounds (Rébeillé *et al.*, 2006). 5,10-Methylenetetrahydrofolate can lose its methylene group to form formaldehyde and tetrahydrofolate, the latter being oxidized as mentioned above. In animal cells, ferritin is able to break the bond between *p*-ABA and the glutamate tail. Since ferritin is present in plastids, it is not unlikely that this process also occurs in plants (Suh *et al.*, 2000). Other possible degradation reactions are a deamination of the pteridine ring (Kobayashi *et al.*, 1967), hydrolysis of the glutamate tail (resulting in a monoglutamate) (Basset *et al.*, 2005) and cleavage of the bond between *p*-ABA and the glutamate tail.
Post-harvest folate degradation in vegetables and fruits can be considerable and therefore nutritionally relevant (Scott *et al.*, 2000). Therefore, it is of paramount importance to comprehend folate stability and explore different strategies to enhance it in order to further improve folate content in crops.

Different approaches can be followed to increase folate stability in plants: (1) engineering towards a more stable folate derivative, (2) association to folate binding proteins, (3) production of anti-oxidants and (4) folate salvage.
As mentioned previously, 5-formyltetrahydrofolate is the most stable natural form of folate. Surprisingly, no function in C1 metabolism has been assigned to this molecule so far. Most possibly, it is an important folate storage form or it may have a regulatory function (Rébeillé *et al.*, 2006). The only enzyme known to use 5-formyltetrahydrofolate is 5-formyltetrahydrofolate cycloligase (5-FCL), which converts the former to 5,10-methenyltetrahydrofolate in an ATP dependent reaction and has a mitochondrial localization in plants (Roje *et al.*, 2002). As mentioned above, knocking out this activity led to 5-formyltetrahydrofolate accumulation and an overall increase of folates by approximately 2-fold, but was associated with impairment of photorespiration and growth retardation in *Arabidopsis* (Goyer *et al.*, 2005). Complexation of folate with folate binding proteins (FBPs) enhances its stability (Hutchinson *et al.*, 2000; Jones and Nixon, 2002). Therefore, overexpression of mammalian FBPs in plants has been proposed (Storozhenko *et al.*, 2005). Additionally, folate binding might create a folate sink, which, in turn, might stimulate folate biosynthesis and/or eliminate possible negative feedback regulation (Rébeillé *et al.*, 2006). Enhancement of antioxidant levels may also be an option to protect plant folates from oxidation. A few attempts have been made to engineer ascorbate levels (Jain and Nessler, 2000; Lorence *et al.*, 2004). However, enhancing the level of one antioxidant could have dissapointing outcomes, since antioxidants are known to act as a network (Blomhoff, 2005). Nevertheless, this approach remains unexplored to date.

Folate salvage

Although folate breakdown products are excreted in animals (Scott, 1984), their destiny in plants remained unclear until recently. Various experiments demonstrated the existence of folate salvage in plants (Orsomando *et al.*, 2006). First of all, the breakdown-products do not accumulate in *Arabidopsis*, tomato and pea, not even under conditions of net folate degradation, when folate synthesis was blocked by sulfanilamide, an inhibitor of DHPS. Second, *Arabidopsis* and tomato tissue are able to convert externally supplied folate degradation products to folate biosynthesis intermediates. Third, the plant folate salvage reactions were detected *in vitro*. Up until now, three reactions are known to recycle folates, two of which are implicated in *p*-ABA salvage, the third one being involved in recycling the pterin moiety (Figure 5).

Side chain reduction of a folate breakdown product, dihydropterin-6-aldehyde (DHPA), produces 6-hydroxymethyldihydropterin (HMDHP), a folate biosynthesis intermediate. Until recently, no pterin aldehyde reductase (PTAR) gene was identified in plants. Experiments in *Arabidopsis*, pea leaves and tomato fruit revealed NADPH-dependent PTAR activity, most possibly resulting from enzymes encoded by the short-chain dehydrogenase/reductase (SDR) gene family (Noiriel *et al.*, 2007b). A candidate *Arabidopsis* PTAR gene has been identified (Noiriel *et al*., 2007b), but another study proved its importance in the modulation of fatty acid composition in seeds (Ajjawi *et al*., 2010). In addition to side chain oxidation of pterins (i.e. DHPA, HMDHP and DHP), ring oxidation also occurs, bearing pterin-6-aldehyde (PA), 6-hydroxymethylpterin (HMP) and pteroate respectively. *Leishmania* and other trypanosomatid parasites have the ability to reduce the ring of these aromatic pterins, a trait that is not shared with plants or *E. coli* (Noiriel *et al.*, 2007a). However, this shortcoming is greatly compensated by high PTAR activity, preventing this kind of oxidation by fast reduction of the pterin side chain (Noiriel *et al.*, 2007b). Nevertheless, the fully oxidized form pterin-6-carboxylate (PtCOOH) cannot be salvaged by bacteria, *Leishmania* or plants (Shiota, 1959; Nare *et al.*, 1997; Stakhov *et al.*, 2002). Consequently, this end-product is a major pterin in dry seeds and harvested fruit (Kohashi *et al*., 1980a; Kohashi *et al.,* 1980b; Diaz de la Garza *et al.*, 2004).

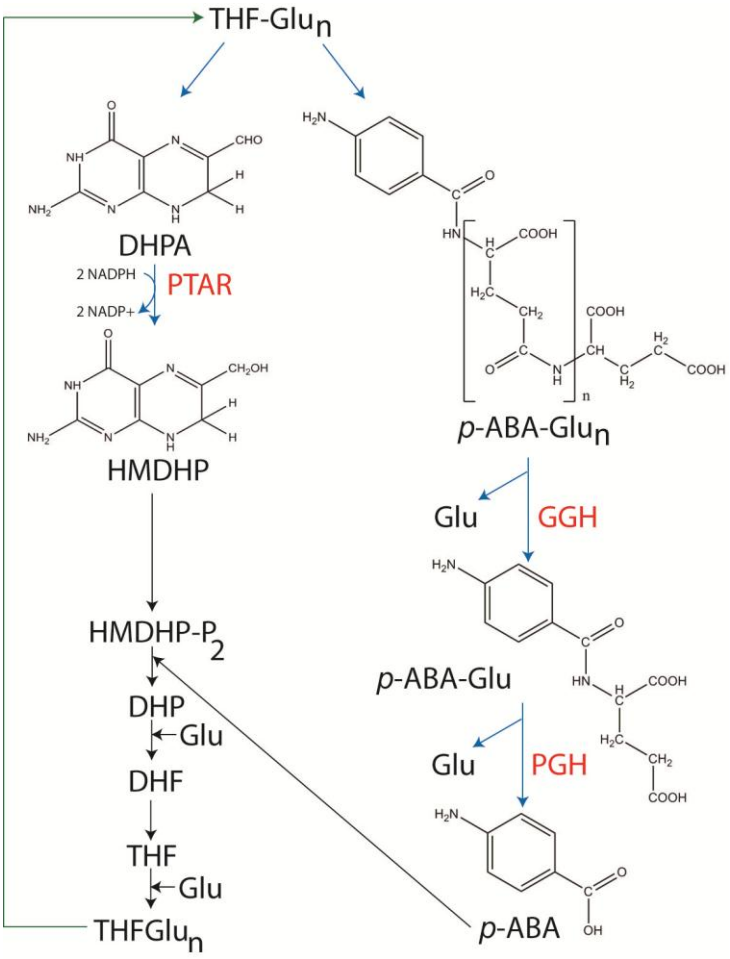


Figure 5. Schematic representation of folate salvage reactions in plants. Tetrahydrofolate polyglutamate degrades in dihydropterin-6-aldehyde and *p*-ABA polyglutamate by oxidation. In plants, up to date, three salvage reactions are known (blue arrows). *p*-ABA and 6-hydroxymethyldihydropterin are able to re-enter the folate biosynthesis cycle (black arrows). In red: folate salvage enzymes. Abbreviations: THF-Glu_n, tetrahydrofolate polyglutamate; DHPA, dihydropterin-6-aldehyde; HMDHP, 6-hydroxymethyldihydropterin; *p*-ABA-Glu_n, para-aminobenzoate polyglutamate; *p*-ABA-Glu, *para*-aminobenzoate monoglutamate; *p*-ABA, *para*-aminobenzoate; HMPHP-P_2, 6-hydroxymethyldihydropterin pyrophosphate; DHP, dihydropteroate; DHF, dihydrofolate; THF, tetrahydrofolate; PTAR, pterin aldehyde reductase; GGH, *gamma*-glutamyl hydrolase; PGH, *p*-ABA-Glu hydrolase.

The folate degradation product *para*-aminobenzoyl polyglutamate (*p*-ABA-Glu_n) is likely able to re-enter the biosynthesis pathway after conversion to *p*-ABA. Therefore, the removal of the polyglutamate tail, an action which is established in two steps, is necessary.

During the first step, *para*-aminobenzoyl monoglutamate (*p*-ABA-Glu_1) is formed. This hydrolysis is mediated by γ-glutamyl hydrolases (GGH), enzymes which are found in mammals as well as in plants (Lin *et al.*, 1993; Suh *et al.*, 2001). These proteins also use folylpolyglutamates as a substrate, as proven by a massive deglutamylation upon disruption of cellular compartmentalization (Melse-Boonstra *et al*., 2002). All plants seem to have one or few GGH genes. A BLAST search revealed 1 gene in rice and maize, two in soybean and cotton and three in *Arabidopsis* (Orsomando *et al*., 2005). Recently, three GGH genes in tomato have been identified (Akthar *et al.*, 2008). Depending on the co-expression of these genes, the gene products appear to have the ability to form heterodimers, as well as homodimers. In *Arabidopsis*, GGH genes are tandemly arranged on chromosome one (Orsomando *et al.*, 2005). The gene products (AtGGH1, AtGGH2 and AtGGH3, respectively) show approximately 35% identity to human GGH. Similar to the latter as well, *Arabidopsis* GGH appear to form dimers (Orsomando *et al.*, 2005). Although they act as exo- or endopeptidases in mammals (McGuire and Coward, 1984), the plant enzymes mainly have endopeptidase activity (Orsomando *et al.*, 2005). Despite the high similarity between the *Arabidopsis* GGH enzymes (70-80% identity), their hydrolysis products are different. AtGGH1 cleaves pentaglutamates predominantly to di- and

triglutamates, whereas AtGGH2 activity yields mainly monoglutamates (Orsomando *et al.*, 2005). Nevertheless, the *Arabidopsis* GGH enzymes appear to complement each other to produce folyl and *p*-ABA monoglutamates. Subcellular fractionation of pea leaves and red beet roots revealed a vacuolar localization of these proteins, where they paradoxically co-occur with folylpolyglutamates (Orsomando *et al.*, 2005). The latter somehow seem to be protected from hydrolysis by the former, most possibly by complexation with folate-binding proteins (Orsomando *et al.*, 2005). The existence of these proteins in plants, however, remains unproven to date.

The folate biosynthesis enzyme dihydropteroate synthase (DHPS) can use *p*-ABA-Glu_1, but not *p*-ABA-Glu_n, as a substrate, yielding dihydrofolate (DHF). However, experiments in *E. coli* and *Plasmodium* showed that the preference to use *p*-ABA instead of *p*-ABA-Glu_1 is 100-fold higher (Ferone, 1973; Swedberg *et al.*, 1979). This preference is less pronounced in *Pneumococcus* (Ortiz, 1970) while being absent in *Lactobacillus* (Shiota *et al.*, 1969). The *in vivo* significance of this reaction in plants is not clear.

The second step in *p*-ABA salvage comprises the hydrolysis of the remaining glutamate from *p*-ABA-Glu_1 to produce the folate biosynthesis intermediate *p*-ABA. This reaction is catalyzed by *para*-aminobenzoylglutamate hydrolase (PGH), an enzyme which is also able to attack folates (Bozzo *et al.*, 2008). This activity is found in pea and *Arabidopsis*, being higher in roots than in leaves, as well as in tomato fruits (Orsomando *et al.*, 2006; Bozzo *et al.*, 2008). Experiments showed the existence of two or more PGH isoforms in plants, of which at least one is a Mn^{2+} containing metalloenzyme (Bozzo *et al.*, 2008). Although the prevalence of PGH proteins is very low, subcellular fractionation revealed a cytosolic/vacuolar and mitochondrial localization of the enzyme activity, with the highest share in the former fraction (89%) (Bozzo *et al.*, 2008).

Folate salvage enzymes have been characterized in much detail in other organisms. The ability to convert *p*-ABA-Glu to *p*-ABA has been demonstrated in bacteria (Hussein *et al.*, 1998; Sherwood *et al.*, 1985) and *Leishmania*, and bacterial cell extracts have the capacity to use pterin-6-aldehyde in folate biosynthesis by reducing this breakdown-product to HMDHP (Shiota, 1959; Brown *et al.*, 1961; Nare *et al.*, 1997). Furthermore, the enzyme activities catalyzing these reductions are present in *Escherichia coli* and *Leishmania* (Mitsuda and Suzuki, 1971; Bello *et al.*, 1994). In *Leishmania major*, a pteridine auxotroph (Ouellette *et al.*, 2002), several experiments were carried-out in relation to pteridine reductase 1 (PTR1). This protein displays high homology with mammalian dihydropterin reductase (DHPR) enzymes, but lacks the corresponding activity (Bello *et al.*, 1994). Instead, it possesses a broad pteridine and folate reductase activity and is capable of reducing both unconjugated and conjugated pterins with an oxidized or dihydro state (Bello *et al.*, 1994), the highest activity being found with the most oxidized forms (Wang *et al.*, 1997). Even though PTR1 is inhibited by MTX, overexpression of the gene ensures MTX resistance, whereas the *ptr1*$^-$ mutant is hypersensitive to this antifolate drug (Bello *et al.*, 1994). A study in *Leishmania tarentolae* gave similar results (Wang *et al.*, 1997).

Folate biofortification

Up until now, the only approach leading to substantial enhancement of folate levels through metabolic engineering is the overexpression of folate biosynthetic enzymes. First attempts to increase folate content in plants concerned the overexpression of GTP cyclohydrolase I (GTPCHI), the first enzyme in pteridine biosynthesis, in *Arabidopsis*, tomato, lettuce and white corn (Hossain *et al.*,

2004; Diaz de la Garza *et al.*, 2004; *Nunes et al*., 2009; Naqvi *et al*., 2009). In tomato, a synthetic gene, based on mouse GTPCHI was inserted. This transgene was under the control of the tomato E8 fruit-specific promoter (Diaz de la Garza *et al.*, 2004). In mammals, GTPCHI is inhibited by the end-product of the pterin biosynthesis pathway, tetrahydrobiopterin. This inhibition is mediated by a GTP cyclohydrolase I feedback regulatory protein (GFRP) (Yoneyama and Hatakeyama, 1998). However, plants lack a GFRP homolog (Basset *et al.*, 2002; Kohashi *et al.*, 1980c) as well as tetrahydrobiopterin; therefore negative feedback inhibition of the mammalian enzyme in plant tissue is unlikely. This engineering approach resulted in a 3- to 140-fold increase of pteridines in tomato fruit, compared with controls, whereas a modest 2-fold average increase of folate content was measured (2.99 nmol/g fresh weight compared with 0.8-2.3 nmol/g fresh weight in controls). The highest increase in folate levels was present in transgenic fruit with a moderate pteridine enhancement (25 nmol/g fresh weight). The accumulating pteridines were neopterin, monapterin and hydroxymethylpterin, as well as their reduced forms and unknown pteridine glycosides. The latter were previously not found in plants, but their existence is known in cyanobacteria and other prokaryotes (Forrest and Van Baalen, 1970; Lin and White, 1988). This high production of pteridines also proved that the enzymes, downstream in pteridine biosynthesis (See paragraph Folate biosynthesis), could handle the increased flux. Concordant with the control plants, the main folates found in transgenic tomato were 5-methyltetrahydrofolate polyglutamate and 5,10-methenyltetrahydrofolate polyglutamate. Measurements of *p*-ABA revealed a total depletion in the GTPCHI overexpressing lines, approximately 90-97% of the *p*-ABA pool was incorporated in folates. Exogenous *p*-ABA supplementation further increased folate content 2,5-10 fold, as compared to controls (Diaz de la Garza *et al*., 2004).

Similar attempts were made in *Arabidopsis* (Hossain *et al.*, 2004) and more recently in lettuce (Nunes *et al*., 2009). For the same reason as mentioned above, a bacterial GTPCHI gene, naturally lacking the negative feedback regulation (*folE* gene from *Escherichia coli*), was used in *Arabidopsis* transformation. Here, pteridine content in leaves raised 1250 fold, in contrast with the folate level, which was a mere two to four times higher than in controls. In lettuce, a synthetic codon-optimized *gchI* gene was inserted, based on the native chicken gene (Nunes *et al*., 2009). Folate content in leaves were 2.1 to 8.5 fold higher than in the corresponding controls, but the pteridine levels are not reported. However, based on the abovementioned reports on tomato and *Arabidopsis* one can predict high levels of pterin accumulation.

In a later round of engineering, tomato was transformed with an *Arabidopsis* cDNA encoding ADC synthase, the enzyme which mediates the first step in *p*-ABA biosynthesis (Diaz de la Garza *et al.*, 2007). This transgene was under the control of the fruit-specific E8 promoter. *P*-ABA levels rose 19-fold, compared with the control and 85% of the produced *p*-ABA was present as its glucose-ester. On the other hand, folate levels did not change.

Overall, these findings clearly indicate that a simultaneous overexpression of pteridine and *p*-ABA biosynthetic genes is necessary to substantially increase folate content. Recently, two groups independently achieved major enhancement of folates in crop plants, respectively in tomato (Diaz de la Garza *et al.*, 2007) and in rice (Storozhenko *et al.*, 2007b) by overexpression of ADC synthase and GTPCHI encoding transgenes. In tomato, the above mentioned ADC synthase overexpressing lines were crossed with the previously obtained GTPCHI overexpressors (Diaz de la Garza *et al.*, 2007). In the resulting transgenic tomato fruits, folate levels were on average 19-fold higher than in controls. However, pteridine and *p*-ABA levels were also very high (20-fold higher than in controls: 50-80 nmol *p*-ABA/g fresh weight and 15-23 nmol pteridines/g fresh weight), pointing to another constraint in

folate biosynthesis flux. Since neither dihydropteroate, nor dihydrofolate or dihydropterin pyrophosphate accumulated in these lines, pteridine import into the mitochondria and/or HPPK activity presumably are the limiting steps. The produced folates were also less extensively polyglutamylated than in controls, possibly because of insufficient FPGS activity. Despite the success of this engineering attempt, it is unclear whether the obtained high pteridine and *p*-ABA levels are nutritionally harmless. *p*-ABA is generally regarded as safe (GRAS) when less than 30 mg/day is consumed. Little is known about pteridine toxicity, although humans can synthesize and catabolize pteridines (Witter *et al.*, 1996). In parallel, transgenic rice lines were obtained, expressing *Arabidopsis* GTPCHI and ADCS transgenes in seeds from a single T-DNA locus (GA-lines) or either one of these genes (G- and A-lines, respectively) (Storozhenko *et al.*, 2007b). Folates accumulated up to 100-fold the level of empty-vector controls in the GA-lines, providing four times the adult RDI in just 100 g of polished raw grains. 5-Methyltetrahydrofolate was the most abundant folate form (89%) in transgenic rice seeds and only 2.6-14% of folates were polyglutamylated, whereas 50% of wild type folates are polyglutamates. Pteridine and *p*-ABA contents rose modestly up to respectively 4-fold and 25-fold the level of empty-vector controls. G-lines showed no folate enhancement, but a 25-fold increase of pteridine content. On the other hand, *p*-ABA content in A-lines was 49-fold higher and, surprisingly, folate levels were 6-fold lower as compared to the controls (Storozhenko *et al.*, 2007b). This is suggested to result from an inhibition of folate biosynthesis by high *p*-ABA concentrations. The ratio of folates/*p*-ABA/pteridines is 1/0.5/0.013 in GA rice lines, while a 1/2.5/0.75 ratio was found in tomato, clearly indicating a more efficient biosynthetic flux in the bio-fortified rice.
Microarray analysis in biofortified tomato proved that the strong enhancement of folate genes influenced less than 20 transcripts more than two-fold, out of 11,000 surveyed. Only 3 genes downstream in the folate biosynthesis pathway were affected (ADCL, DHNA and mitochondrial FPGS), indicating feedforward control by pathway intermediates (Waller *et al.*, 2009).

In bacteria, similar engineering approaches were used to boost folate production. In *Lactococcus lactis*, a gene cluster with five genes (*folA*, *folB*, *folKE*, *folP* and *folC* encoding respectively DHFR, DHNA, HPPK/GTPCHI, DHPS and DHFS/FPGS proteins) is present in the genome (Sybesma *et al.*, 2003a). Overexpression of *folKE* resulted in a 10-fold accumulation of extracellular folates, whereas total folate content rose 3-fold. Co-overexpression of *folC* (DHFS/FPGS) retained the folates in the cells, proving that the obtained extracellular folate enhancement is due to secretion of folate monoglutamates. Experiments with GGH from rat and human origin strengthened this theory (Sybesma *et al.*, 2003b). When *folA* (DHFR) was overexpressed, folate content dropped, pointing to a feedback inhibition imposed by reduced folates on folate biosynthesis in *Lactococcus*. Other relevant work on improving microbial folate biosynthesis has been done in *Bacillus subtilis* (Zhu *et al.*, 2005) and *Streptococcus thermophilus* (Lin and Young, 2000). Overexpression of complete folate biosynthesis gene clusters, a strategy that has been used in *Lactobacillus gasseri* (Wegkamp *et al.*, 2004), *Lactococcus lactis* (Wegkamp *et al.*, 2007) and *Lactobacillus reuteri* (Santos *et al.*, 2008) could even increase folate levels 100-fold (Santos *et al.*, 2008).

Exploring natural variation in folate levels could be an alternative way to enhance folate content (Bekaert *et al.*, 2008), although this is considered to be limited through classical breeding (Rébeillé *et al.*, 2006). Mapping quantitative trait loci (QTL), integrated in molecular marker-assisted selection, is a way to overcome this problem, but this remains unexplored to date.

Towards 'nutritionally complete' crops?

Although numerous examples of biofortification exist, the majority of them deal with the enhancement of a single nutrient. Since a number of staple crops are poor in vitamins and minerals, simultaneous enhancement of several nutritionally important compounds is desirable. Moreover, since the plant's system attempts to maintain homeostasis, a multi-trait metabolic engineering approach, modulating multiple enzymatic steps at a time, is often necessary to enhance nutritional value of transgenic crops. This results in the need to 'stack' or 'pyramide' traits, which naturally faces the challenge of stable and coordinated expression of multiple (trans)genes, each involved in their own metabolic pathway.

A number of conventional and more recent approaches exist to introduce several genes/traits into plants, all of them having their own advantages and drawbacks (extensively reviewed in Halpin, 2005). Conventional breeding (based on quantitative trait loci, QTL's) appears to be less useful for this purpose as compared to transgenic approaches, because the former relies on alleles already in the species genepool, hence its limitation, at least in those species where few-fold variation is found for a given nutrient. Our recent analyses revealed up to 2-fold difference between the highest and lowest levels of total folate in seeds of 12 greenhouse-grown rice varieties (Figure 6 and Table 2). The amount of total folate ranged from 32 µg/100g fresh weight in Rok, an *Oryza sativa* variety from Sierra Leone, to 68 µg/100g fresh weight in the Blue belle variety from Guyana. This variation mainly originates from the difference in accumulation of 5-methyltetrahydrofolate, which is the major folate form found in rice seeds (Storozhenko *et al.*, 2007b). A similar extent of variation in folate levels was observed for 130 winter and 20 spring wheat genotypes, ranging from 364 to 774 ng/g of dry matter in winter wheat and from 323 to 741 ng/g of dry matter in spring wheat varieties (Piironen *et al*., 2008). Thus, even though there is a clear potential to exploit natural variation, it may be hard to reach the level of enhancement needed to alleviate folate deficiency. Moreover, identifying useful QTL's and introgression into elite cultivars is a complex and time-consuming task for a single vitamin or mineral, making this method impracticable for multiple traits. Alternatively, transgenic approaches through direct (e.g. biolistics) or indirect (e.g. *Agrobacterium* mediated) transfer of genes are proven to be more convenient in creating these 'supercrops'. Nevertheless, several restrictions still limit its success.

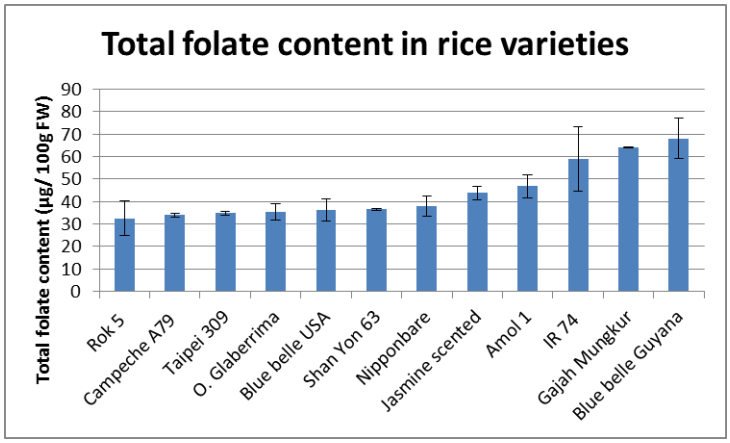


Figure 6. Total folate levels (in µg/100g fresh weight) in seeds of different rice varieties. Error bars indicate standard deviations (n=2).

Iterative strategies, such as crossing and re-transformation, can be valuable but are labour-intensive, time demanding and the introduced traits are prone to segregation. Typically, novel traits are introduced by crossing or transforming homozygote parents but finding the latter alone can take two or more generations. Moreover, re-transformation requires the use of different selectable marker genes. Luckily, several strategies have been developed to enable the removal of those markers (reviewed in Hohn *et al*., 2001; Hare and Chua, 2002).

Co-transformation, either directly or indirectly, is a quick and simple strategy for the introduction of multiple genes/traits into plants. However, it is very likely that multiple copies of each expression cassette are introduced in the plant, especially when using biolistic transformation methods (Maqbool and Christou, 1999). On the other hand, co-transformation is considered to be an effective method to limit the number of selectable markers (Halpin, 2005).

Table 2. Individual and total folate concentrations (in µg/100g fresh weight) in seeds of different rice varieties. RSD: Relative standard deviation (%); THF: tetrahydrofolate; 5MeTHF: 5-Methyltetrahydrofolate; FA: Folic acid; 5FoTHF: 5-Formyltetrahydrofolate; 10FoFA: 10-Formylfolic acid; 5,10-CH^{+}THF: 5,10-Methenyltetrahydrofolate; nq: not quantifiable.

	THF		5MeTHF		FA		5FoTHF		10FoFA		5,10 CH+ THF		Sum
	(µg/100g)	% RSD	(µg/100g)	% RSD	(µg/100g)	% RSD	(µg/100g)	% RSD	(µg/100g)	% RSD	(µg/100g)	% RSD	(µg/100g)
Rok 5	1.54	38.3	21.06	29.9	0.75	25.1	4.90	1.3	2.77	18.6	1.44	9.6	32.45
Campeche A79	1.12	0.2	21.73	7.3	nq		5.67	35.6	4.78	24.8	0.63	54.2	33.92
Taipei 309	1.98	6.0	19.13	28.2	nq		9.94	5.0	2.83	13.0	0.88	13.2	34.77
O. Glaberrima	2.01	20.9	19.71	3.8	0.98	16.4	6.02	37.8	5.20	1.7	1.47	20.6	35.39
Blue belle USA	1.30	11.1	26.43	13.3	0.15	0.2	4.94	35.1	2.52	4.9	0.76	8.5	36.10
Shan Yon 63	2.24	7.0	21.76	7.1	nq		5.60	12.1	5.42	14.8	1.49	7.8	36.52
Nipponbare	1.50	14.7	28.61	17.0	0.24	134.9	4.41	14.5	2.38	11.1	0.92	1.7	38.07
Jasmine scented	2.04	17.9	30.21	1.6	0.81		5.76	32.3	4.03	9.2	0.99	27.8	43.85
Amol 1	2.21	13.8	34.70	15.7	0.03		5.53	2.4	2.91	11.5	1.44	5.4	46.82
IR 74	1.50	1.6	45.42	24.2	0.49	66.1	4.79	30.1	5.95	19.3	0.84	42.9	59.00
Gajah Mungkur	1.87	10.5	52.69	3.9	nq		5.08	20.9	3.14	33.0	1.25	4.5	64.04
Blue belle Guyana	1.62	32.9	50.62	17.9	0.81	44.8	10.28	11.1	3.52	11.7	1.30	11.5	68.15

Besides stable integration, a coordinated expression of the introduced genes is crucial. Linked genes do not necessarily have similar expression levels, even if they are controlled by the same promoters (Peach and Velten, 1991). Moreover, the chance that transgenes residing at independent genomic loci are equally expressed is even smaller. This can be caused by position effects and gene silencing mechanisms (Halpin, 2005).

The maximum number of transgenes/traits that can be introduced in a crop, if there is a limit at all, remains to be discovered. The co-transformation strategy has already resulted in the insertion of 14 different genes into rice plants (Chen *et al*., 1998). Recently, a multivitamin South African white corn has been produced by engineering three independent metabolic pathways (Naqvi *et al*., 2009). This transgenic corn has an enhanced content of three vitamins: β-carotene (169-fold), ascorbate (6-fold) and folate (2-fold). This predictably modest elevation of folate levels was generated through the introduction of the *E. coli folE* gene, controlled by the barley D-hordein promoter. The production of 'nutritional supercrops' is at present in its infancy, and regulatory approval is probably the biggest hurdle to be taken.

Case-study: ex-ante evaluation of folate biofortified rice in Shanxi Province, China

Shanxi province, a poor rural region in North China, is characterized by one of the highest NTD prevalence rates in the world (Li *et al.,* 2006; Gu *et al.,* 2007). This high NTD risk is associated with a low intake of folic acid (Ren *et al.,* 2006; Li *et al.,* 2007). Furthermore, these folic acid supplements are often only available in cities, which makes Shanxi province an excellent region to study the introduction of folate biofortification. Rice, initially low in folate content, was selected as the food vehicle for biofortification as it is one of the most important staple crops in Shanxi province.
Figure 7 presents the conceptual framework to evaluate the potential of folate biofortified rice and consists of two parts. The first part explores consumer's acceptance of folate biofortified rice. This acceptance rate is the precondition for the commercialization of folate biofortified rice in Shanxi province. Based upon scientific evidence, this part contains the most important consumer characteristics that influence acceptance of a Genetically Modified (GM) food product, namely socio-demographic indicators, knowledge and consumer perceptions of GM food (Bonti-Ankomah and Yiridoe, 2006; Costa-Font *et al.,* 2008). The second part indicates the well-established relationship between folate intake and NTD prevalence, also known as the dose-response relationship (Moore *et al.,* 2003). This relationship measures the impact of a folate biofortified diet on the main adverse health outcome of folate deficiency. The folate intake depends on the consumer's acceptance of folate biofortified rice.

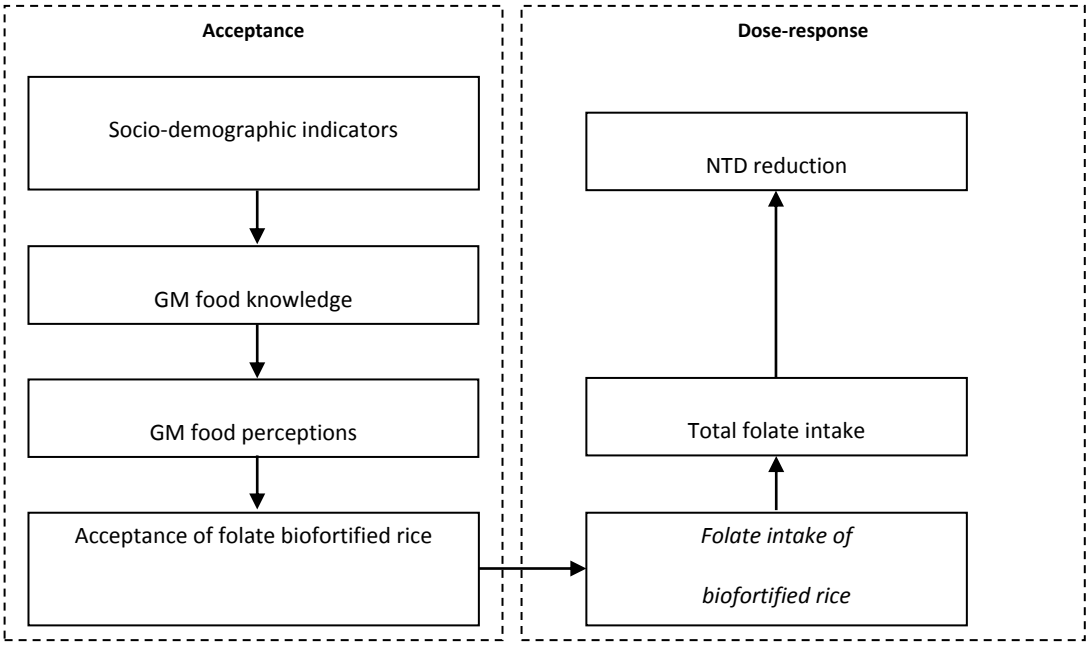


Figure 7. Conceptual framework of the impact of folate biofortified rice on the reduction of neural tube defects.

The consumer survey on the determinants of biofortified rice acceptance encompasses random face-to-face interviews with 944 rice consumers from Shanxi province, China (De Steur *et al.*, 2010). Regarding subjective or perceived knowledge of GM food, 47.8 % of the consumers believe to know what GM food is. Subjective knowledge in rural regions is significantly lower than urban regions in

China (Huang *et al.,* 2006). Objective GM food knowledge, measured by six reliable true-or-false statements, is even lower, with a mean score of 39.2 %. This reveals misjudgement of subjective GM food knowledge (Li *et al.,* 2002).

Consumer perceptions toward GM food are categorized into four groups of statements: benefits, risks, safety and price. After reliability analysis ($\alpha>0.6$), means are analysed for each group. In general, consumers evaluate the four categories as slightly positive, but perceptions on safety of GM food are significantly more positive (Paired Sample t-test, $p<0.001$). Regarding folate biofortified rice, consumers evaluate the health benefits of GM food as positive, while the potential of being a medical substitute is considered as neutral.

Women are considered as the target group of folate biofortified rice consumption. In general, 55.4 % of all female rice consumers accept this biofortified rice, while 32.3 % react indifferent and 12.3 % reject this product. These promising findings are in line with results from recent Chinese studies (Li *et al.,* 2002; Ho *et al.,* 2006; Huang *et al.,* 2006).

The acceptance rate (55.4 %) is considered as the coverage rate of the biofortified rice intervention. It determines the folate intake (dose) and, therefore, the number of NTDs that can be prevented by women's biofortified diet (response).

Research on the dose-response relationship in Shanxi province has shown that 79 % to 85 % of all cases of NTDs can be prevented by consumption of the RDI of 400 µg folates (Berry *et al.,* 1999). In other words, if a daily folate intake below 400 µg is considered as folate deficiency, between 79 % and 85 % of all NTDs in Shanxi province are thought to be attributable to folate deficiency. Worldwide, the impact of a daily folate intake of 400 µg is significantly lower, with a reduction of 50 % to 70 % of all NTDs (Daly *et al.,* 1997; Molloy and Scott, 2001).

Based on the level of folates in biofortified rice after processing and taking into account losses due to bioavailability (Storozhenko *et al.,* 2007b), the current rice consumption, and folate intake in Shanxi province (Zhang *et al.,* 2008; CNGOIC, 2009), the total folate intake under biofortification amounts to 503.7 µg per day, per person. This dose is significantly higher than that by which the risk of having a baby with an NTD caused by folate deficiency is largely prevented (Berry *et al*., 1999; see above).

Dependent on the contribution of folate deficiency to NTDs (79% to 85 %) and the coverage rate (55.4 %), women's folate biofortified diet can prevent between 3443 and 3705 births from an NTD (Table 3). This is a theoretical reduction of 44 % to 47 % in the total burden of NTDs.

Table 3. Impact of folate biofortified rice on the reduction of neural tube defects in Shanxi province.

Numbers for Shanxi	**Contribution of folate deficiency to NTDs***		**Source**
	79 %	**85 %**	
Births	394674	394674	(Shanxi Province Statistical Bureau, 2007)
NTD prevalence rate	199.38	199.38	per 10000 births in 2004 (Gu *et al.*, 2007)
NTDs	7869.01	7869.01	
NTDs < folate deficiency	6216.52	6688.66	
Coverage rate	55.40 %	55.40 %	
Average # women recovering	300.00	322.78	
Average # NTDs	3443.95	3705.52	

*Based on Berry *et al*., (1999)

Concluding remarks

Despite successful attempts to create folate biofortified foods, a lot of work still needs to be done in order to effectively fight folate deficiency in developing countries. A thorough knowledge and understanding of folate stability and bioavailability could provide new insights to help eradication of this global problem.

Acknowledgments

This work was supported by a grant from Ghent University (Bijzonder Onderzoeksfonds, GOA 1251204) to D.V.D.S., W.L. and J.V.
D.B. and V.D.B. are indebted to FWO (Fonds voor Wetenschappelijk Onderzoek-Vlaanderen) and IWT (Instituut voor de Aanmoediging van Innovatie door Wetenschap en Technologie) respectively, for a PhD grant.

References

Ajjawi I, Lu Y, Savage LJ, Shannon MB, Last RL. 2010. Large-scale reverse genetics in Arabidopsis: Case studies from the chloroplast 2010 project. Plant Physiology **152**, 529-540.

Akhtar TA, McQuinn RP, Naponelli V, Gregory JF III, Giovannoni JJ, Hanson AD. 2008. Tomato gamma-glutamylhydrolases: expression, characterization, and evidence for heterodimer formation. Plant Physiology **148**, 775-785.

Allfrey VG, Faulkner R, Mirsky AE. 1964. Acetylation and methylation of histones and their possible role in the regulation of RNA synthesis. Proceedings of the National Academy of Sciences of the United States of America **51**, 786-794.

Appling DR. 1991. Compartmentation of folate-mediated one-carbon metabolism in eukaryotes. FASEB Journal **5**, 2645-2651.

Bailey LB. 2004. Folate and Vitamin B12 recommended intakes and status in the United States. Nutrition Reviews **62**,S14-S20.

Basset G, Quinlivan EP, Ziemak MJ, Diaz de la Garza R, Fischer M, Schiffmann S, Bacher A, Gregory JF III, Hanson AD. 2002. Folate synthesis in plants: the first step of the pterin branch is mediated by a unique bimodular GTP cyclohydrolase I. Proceedings of the National Academy of Sciences of the United States of America **99**, 12489-12494.

Basset GJ, Quinlivan EP, Ravanel S, Rébeillé F, Nichols BP, Shinozaki K, Seki M, Adams-Philips LC, Giovannoni JJ, Gregory JF III, Hanson AD. 2004a. Folate synthesis in plants: the *p*-aminobenzoate branch is initiated by a bifunctional PabA-PabB protein that is targeted to plastids. Proceedings of the National Academy of Sciences of the United States of America **101**, 1496-1501.

Basset GJC, Ravanel S, Quinlivan EP, White R, Giovannoni JJ, Rébeillé F, Nichols BP, Shinozaki K, Seki M, Gregory JF III, Hanson AD. 2004b. Folate synthesis in plants: the last step of the *p*-aminobenzoate branch is catalyzed by a plastidial aminodeoxychorismate lyase. Plant Journal **40**, 453-461.

Basset GJC, Quinlivan EP, Gregory JF III, Hanson AD. 2005. Folate synthesis and metabolism in plants and prospects for biofortification. Crop Science **45**, 449-453.

Bedhomme M, Hoffmann M, McCarthy EA, Gambonnet B, Moran RG, Rébeillé F, Ravanel S. 2005. Folate metabolism in plants: an *Arabidopsis* homolog of the mammalian mitochondrial folate transporter mediates folate import into chloroplasts. The Journal of Biological Chemistry **280**, 34823-34831.

Bekaert S, Storozhenko S, Mehrsahi P, Bennet MJ, Lambert W, Gregory JF III, Schubert K, Hugenholtz J, Van Der Straeten D, Hanson AD. 2008. Folate biofortification in food plants. Trends in Plant Science **13**, 28-34.

Bello AR, Nare B, Freedman D, Hardy L, Beverley SM. 1994. PTR1: a reductase mediating salvage of oxidized pteridines and methotrexate resistance in the protozoan parasite *Leishmania major*. Proceedings of the National Academy of Sciences of the United States of America **91**, 11442-11446.

Berry R, Li Z, Erickson D, Li S, Moore C, Wang H, Mulinare J, Zhao P, Wong L-Y , Gindler J, Hong S-X, Correa A. 1999. Prevention of Neural-Tube Defects with folic acid in China. The New England Journal of Medecine **341**, 1485-1491.

Berry RJ, Li Z. 2002. Folic Acid Alone Prevents Neural Tube Defects: Evidence from the China Study. [Letter]. Epidemiology **13**, 114-116.

Besson V, Rébeillé F, Neuburger M, Douce R, Cossins EA. 1993. Effects of tetrahydrofolate polyglutamates on the kinetic parameter of serine hydroxymethyltransferase and glycine decarboxylase from pea leaf mitochondria. Biochemical Journal **292**, 425-430.

Bewley JD. 1997. Seed germination and dormancy. Plant Cell **9**, 1055-1066.

Blackley RL. 1969. The biochemistry of folic acid and related pteridines. **In**: *Frontiers of Biology, Vol. 13.* Neuberger, A. and Tatum, E.L., Eds., North Holland, Amsterdam and London.

Block MA, Tewari AK, Albrieux C, Marechal E, Joyard J. 2002. The plant S-adenosyl-L-methionine: Mg-protoporphyrin IX methyltransferase is located in both envelope and thylakoid chloroplast membranes. European Journal of Biochemistry **269**, 240-248.

Blomhoff R. 2005. Dietary antioxidants and cardiovascular disease. Current Opinion in Lipidology **16**, 47-54.

Bognar AL, Osborne C, Shane B, Singer SC, Ferone R. 1985. Folylpoly-gamma-glutamate synthetase-dihydrofolate synthetase-cloning and high expression of the *Escherichia coli* FolC gene and purification and properties of the gene product. Journal of Biological Chemistry **260**, 5625-5630.

Bonaa KH, Njolstad I, Ueland PM, Schirmer H, Tverdal A, Steigen T, Wang H, Nordrehaug JE, Arnesen E, Rasmussen K. 2006. Homocysteine lowering and cardiovascular events after acute myocardial infarction. The New England Journal of Medecine **354**, 1578-1588.

Bonti-Ankomah S, Yiridoe E. 2006. *Organic and conventional food: a literature review of the economics of consumer perceptions and preferences.* Canada, Organic Agriculture Centre of Canada.

Bozzo GG, Basset GJC, Naponelli V, Noiriel A, Gregory JF III, Hanson AD. 2008. Characterization of the folate salvage enzyme *p*-aminobenzoylglutamate hydrolase in plants. Phytochemistry. **69**, 29-37.

Brown GM, Weisman RA, Molnar DA. 1961. The biosynthesis of folic acid. 1. Substrate and cofactor requirements for enzymatic synthesis by cell-free extracts of *Escherichia coli*. Journal of Biological Chemistry **236**, 2534-2543.

Camara D, Richefeu-Contesto C, Gambonnet B, Dumas R, Rébeillé F. 2011. The synthesis of pABA : Coupling between the glutamine amidotransferase and aminodeoxychorismate synthase domains of the bifunctional aminodeoxychorismate synthase from *Arabidopsis thaliana*. Archives of Biochemistry and Biophysics **505**, 83-90.

Cella R, Parisi B. 1993. Dihydrofolate reductase and thymidylate synthase in plants – an open problem. Physiologia Plantarum **88**, 509-521.

Chen L, Chan SY, Cossins EA. 1997. Distribution of folate derivatives and enzymes for synthesis of 10-formyltetrahydrofolate in cytosolic and mitochondrial fractions of pea leaves. Plant Physiology **115**: 299-309.

Chen LL, Marmey P, Taylor NJ, Brizard JP, Espinoza C, D'Cruz P, Huet H, Zhang SP, de Kochko A, Beachy RN, Fauquet CM. 1998. Expression and inheritance of multiple transgenes in rice plants. Nature Biotechnology **16**, 1060-1064.

Cherian A, Seena S, Bullock RK, Antony AC. 2005. Incidence of neural tube defects in the least-developed area of India: a population-based study. Lancet **366**, 930-931.

Choi SW, Friso S. 2005. Interactions between folate and aging for carcinogenesis. Clinical Chemistry and Laboratory Medecine **43**, 1151-1157.

Choumenkovitch SF, Selhub J, Wilson PWF, Rader JI, Rosenberg IH, Jacques PF. 2002. Folic acid intake from fortification in United States exceeds predictions. Journal of Nutrition **132**, 2792-2798.

CNGOIC. 2009. *Food Grains Weekly Report.* Beijing, National Grain and Oils Information Center.

Cole BF, Baron JA, Sandler RS, Haile RW, Ahnen DJ, Bresalier RS, Mckeown-Eyssen G, Summers RW, Rothstein R, Burke CA, Snover DC, Church TR, Byers T. 2005.A randomised trial of folic acid to prevent colorectal adenomas. Proceedings of the American Association for Cancer Research **46**, 4399.

Cook RJ. 2001. Disruption of histidine catabolism in NEUT2 mice. Archives of Biochemistry and Biophysics **392**, 226-232.

Cossins EA. 2000. The fascinating world of folate and one-carbon metabolism. Canadian Journal of Botany **78**, 691-708.

Costa-Font M, Gil J, Traill W. 2008. Consumer acceptance, valuation of and attitudes towards genetically modified food: Review and implications for food policy. Food Policy **33**, 99-111.

Crosti P, Malerban M, Blanchetti R. 1993. Growth-dependent changes of folate metabolism and biosynthesis in cultured *Daucus carota* cells. Plant Science **88**, 97-106.

Dallas WS, Gowen JE, Ray PH, Cox MJ, Dev IK. 1992. Cloning, sequencing, and enhanced expression of the dihydropteroate synthase gene of *Escherichia coli* MC4100. Journal of Bacteriology **174**, 5961-5970.

Daly S, Mills J, Molloy A, Conley M, Lee Y, Kirke P, Weir D, Scott J. 1997. Minimum effective dose of folic acid for food fortification to prevent neural-tube defects. Lancet **350**, 1666-1669.

Daubner SC, Matthews RG. 1982. Purification and properties of methylenetetrahydrofolate reductase from pig liver. Journal of Biological Chemistry **257**, 140-145.

DeSouza L, Shen Y, Bognar AL. 2000. Disruption of cytoplasmic and mitochondrial folylpolyglutamate synthetase activity in *Saccharomyces cerevisiae*. Archives of Biochemistry and Biophysics **376**, 299-312.

De Steur H, Gellynck X, Storozhenko S, Liqun G, Lambert W, Van Der Straeten D, Viaene J. 2010. Willingness to accept and purchase genetically modified rice with high folate content in Shanxi province, China. Appetite **54**, 118-125.

De Wals P, Tairou F, Van Allen MI, Uh S, Lowry RB, Sibbald B, Evans JA, Van den Hof MC, Zimmer P, Crowley M, Fernandez B, Lee NS, Niyonsenga T. 2007. Reduction in neural-tube defects after folic acid fortification in Canada. The New England Journal of Medecine **357**, 135-142.

Diaz de la Garza R, Quinlivan EP, Klaus SM, Basset GJ, Gregory JF III, Hanson AD. 2004. Folate biofortification in tomatoes by engineering the pteridines branch of folate synthesis. Proceedings of the National Academy of Sciences of the United States of America **101**, 13720-13725.

Diaz de la Garza RI, Gregory JF III, Hanson AD. 2007. Folate biofortification of tomato fruit. Proceedings of the National Academy of Sciences of the United States of America **104**,4218-4222.

Douce R, Bourguignon J, Neuburger M, Rébeillé F. 2001. The glycine decarboxylase system: a fascinating complex. Trends in Plant Science. **6**, 167-176.

Eudes A, Bozzo GG, Waller JC, Naponelli V, Lim EK, Bowles DJ, Gregory JF III, Hanson AD. 2008. Metabolism of the folate precursor *p*-aminobenzoate in plants. Glucose ester formation and vacuolar storage. Journal of Biological Chemistry **283**, 15451-15459.

Fan CL, Brown GM. 1976. Partial purification and properties of guanosine triphosphate cyclohydrolase from *Drosophila melanogaster*. Biochemical Genetics **14**, 259-270.

Ferone R. 1973. The enzymatic synthesis of dihydropteroate and dihydrofolate by *Plasmodium berghei*. The Journal of Protozoology **20**, 459-464.

Finglas PM, de Meer K, Molloy A, Verhoef P, Pietrzik K, Powers HJ, Van Der Straeten D, Jagerstad M, Varela-Moreiras G, van Vliet T, Havenaar R, Buttriss J. 2006. Research goals for folate and related B vitamin in Europe. European Journal of Clinical Nutrition **60**, 287-294.

Food and Drug Administration. 1996. Food standards: amendments of standards of identity for enriched grain products to require addition of folic acid. Final Rule. 21 CFR Parts 136, 137, **139**: 8781-8807.

Forrest HS, Van Baalen C. 1970. Microbiology of unconjugated pteridines. Annual Review of Microbiology **34**, 91-108.

Fowler B. 2001. The folate cycle and disease in humans. Kidney International **78**, S221-S229.

Freemantle SJ, Taylor SM, Krystal G, Moran RG. 1995. Upstream organization of and multiple transcripts from the human folylpoly-gamma-glutamate synthase gene. Journal of Biological Chemistry **270**, 9579-9584.

Gambonnet B, Jabrin S, Ravanel S, Karan M, Douce R, Rébeillé F. 2001. Folate distribution during higher plant development. Journal of the Science of Food and Agriculture **81**, 835-841.

Gao YS, Alvarez C, Nelson DS, Sztul E. 1998. Molecular cloning, characterization, and dynamics of rat formiminotransferase cyclodeaminase, a Golgi-associated 58 kDa protein. Journal of Biological Chemistry **273**, 33825-33834.

Gao YL , Guo F , Gokavi S, Chow A, Sheng QH, Guo MR. 2008. Quantification of water-soluble vitamins in milk-based infant formulae using biosensor-based assays. Food Chemistry **110**, 769-776.

Garcia AA, Day AG, Zanibbi K, Zunzunegui MV. 2008. Long-term effects of folic acid fortification and B-vitamin supplementation on total folate, homocysteine, methylmalonic acid and cobalamin in older adults. Canadian Journal of Public Health **99**, 428-433.

Geisel J. 2003. Folic acid and neural tube defects in pregnancy – A review. The Journal of Perinatal and Neonatal Nursing **17**, 268-279.

Goyer A, Illarionova V, Roje S, Fischer M, Bacher A, Hanson AD. 2004. Folate biosynthesis in higher plants. cDNA cloning, heterologous expression, and characterization of dihydroneopterin aldolases. Plant Physiology **135**, 103-111.

Goyer A, Collakova E, Diaz de le Garza R, Quinlivan EP, Williamson J, Gregory JF III, Shachar-Hill Y, Hanson AD. 2005. 5-Formyltetrahydrofolate is an inhibitory but well tolerated metabolite in *Arabidopsis* leaves. Journal of Biological Chemistry **280**, 26137-26142.

Green JM, Merkel WK, Nichols BP. 1992. Characterization and sequence of *Escherichia coli* PabC, the gene encoding aminodeoxychorismate lyase, a pyridoxal phosphate-containing enzyme. Journal of Bacteriology **174**, 5317-5323.

Green NS. 2002. Folic acid supplementation and prevention of birth defects. Journal of Nutrition **132**, 2356S-2360S.

Gregory JF III. 1989. Chemical and nutritional aspects of folate research: analytical procedures, methods of folate synthesis, stability, and bioavailability of dietary folates. Advances in Food and Nutrition Research **33**, 1-101.

Gregory JF. 1995. The bioavailability of folate. **In**: *Folate in Health and Disease.* pp.195-235. Bailey, B.L., Ed., Marcel Dekker, New York.

Gregory JF. 1997. The bioavailability of vitamin B-6. European Journal of Clinical Nutrition **51**, S43-S48.

Gregory JF III, Quinlivan EP. 2002. In vivo kinetics of folate metabolism. Annual Review of Nutrition **22**, 199-200.

Gu X, Lin L, Zheng X, Zhang T, Song X, Wang J, Li X, Li P, Chen P, Wu J, Wu L., Liu J. 2007. High Prevalence of NTDs in Shanxi Province: A Combined Epidemiological Approach. Birth defects research. Part A, Clinical and Molecular Teratology **79**, 702-707.

Guldener U, Koehler GJ, Haussmann C, Bacher A, Kricke J, Becher D, Hegemann JH. 2004. Characterization of the *Saccharomyces cerevisiae* Fol 1 protein: Starvation for C1 carrier induces pseudohyphal growth. Molecular Biology of the Cell **15**, 3811-3828.

Gutstein S, Bernstein LH, Levy L, Wagner G. 1973. Failure of response to N^5-methyltetrahydrofolate in combined folate and B12 deficiency: Evidence in support of the "folate trap" hypothesis. Digestive Diseases and Sciences **18**, 142-146.

Halpin C. 2005. Gene stacking in transgenic plants – the challenge for the 21st century plant biotechnology. Plant Biotechnology Journal **3**, 141-155.

Hanson AD, Roje S. 2001. One-carbon metabolism in higher plants. Annual Review of Plant Physiology and Plant Molecular Biology **52**, 119-137.

Hanson AD, Gregory JF. 2002. Synthesis and turnover of folates in plants. Current Opinion in Plant Biology **5**, 244-249.

Hare PD, Chua NH. 2002. Excision of selectable marker genes from transgenic plants. Nature Biotechnology **20**, 575-580.

Hennig D, Scales SJ, Moreau A, Murley LL, De Mey J, Kreis TE. 1998. A formiminotransferase cyclodeaminase isoform is localized to the Golgi complex and can mediate interaction of trans-Golgi netwotk-derived vesicles with microtubules. Journal of Biological Chemistry **273**, 19602-19611.

Hillcoat BL, Nixon PF, Blakley RL. 1967. Effects of substrate decomposition on the spectrophotometric assay of dihydrofolate reductase. Analytical Biochemistry **21**, 178-189.

Ho P, Vermeer E, Zhao J. 2006. Biotechnology and food safety in China: consumer's acceptance or resistance? Development and Change **37**, 227-254.

Hohn B, Levy AA, Puchta H. 2001. Elimination of selectin markers from transgenic plants. Current Opinion in Biotechnology **12**, 139-143.

Holland B, Undwin ID, Buss DH. 1996. *Vegetables, Herbs and Spices. Fruit and Nuts. Fifth supplement to McCance and Widdowson's 'The Composition of Foods'.* Cambridge the Royal Society of Chemistry and Ministry of Agriculture Fisheries and Food, Laboratory of the Government Chemist.

Hossain T, Rosenberg I, Selhub J, Kishore G, Beachy R, Schubert K. 2004. Enhancement of folate in plants through metabolic engineering. Proceedings of the National Academy of Sciences of the United States of America **101**, 5158-5163.
Hrazdina G, Jensen RA. 1992. Spatial organization of enzymes in plant metabolic pathways. Annual Review of Plant Physiology and Plant Molecular Biology **43**, 241-267.

Huang J, Qiu H, Bai J, Pray C. 2006. Awareness, acceptance of and willingness to buy genetically modified foods in Urban China. Appetite **46**, 144-151.

Hubner RA, Houlston RD, Muir KR. 2007. Should folic acid fortification be mandatory? No. BRITISH MEDICAL JOURNAL **334**, 1253.

Hussein MJ, Green JM, Nichols BP. 1998. Characterization of mutations that allow *p*-aminobenzoyl-polyglutamate utilization by *Escherichia coli*. Journal of Bacteriology **180**, 6260-6268.

Hutchinson ML, Jones MD, Nixon PF. 2000. Stability of labile folate is enhanced by binding to milk folate binding protein. Australian Journal of Dairy Technology **55**, 98.

Imeson HC, Zheng L, Cossins EA. 1990. Folylpolyglutamate derivates of *Pisum sativum* L. Determination of polyglutamate chain lengths by high-performance liquid chromatography following conversion to p-aminobenzoylpolyglutamates. Plant and Cell Physiology **31**, 223-231.

Jabrin S, Ravanel S, Gambonnet B, Douce R, Rébeillé F. 2003. One-carbon metabolism in plants. Regulation of tetrahydrofolate synthesis during germination and seedling development. Plant Physiology **131**, 1431-1439.

Jacques PF, Selhub J, Bostom AG, Wilson PW, Rosenberg IH. 1999. The effect of folic acid fortification on plasma folate and total homocysteine concentrations. The New England Journal of Medicine **340**, 1449-1454.

Jain AK, Nessler CL. 2000. Metabolic engineering of an alternative pathway for ascorbic acid biosynthesis in plants. Molecular Breeding **6**, 73-78.

Jiang L, Liu YY, Sun H, Han YT, Li JL, Li CK, Guo WZ, Meng HY, Li S, Fan YL, Zhang CY. 2013. The mitochondrial folylpolyglutamate synthetase gene is required for nitrogen utilization during early seedling development in Arabidopsis. Plant Physiology **161**, 971-989.

Johnson JL, Rajagopalan KV. 1985. Pterins in other metabolic pathways. **In**: *Folates and Pterins, Vol. 2.* pp. 383-399. Blackley, R.L. and Benkovic, S., Eds.,J. Wiley (Interscience), New York.

Jones ML, Nixon PF. 2002. Tetrahydrofolate are greatly stabilized by binding to bovine milk folate-binding protein. Journal of Nutrition **132**, 2690-2694.

Kaufman S, Kaufman EE. 1985. Tyrosine hydroxylase. **In**: *Folates and Pterins, Vol. 2.* pp. 251-352. Blackley, R.L. and Benkovic, S., Eds.,J. Wiley (Interscience), New York.

Klaus SM, Wegkamp A, Sybesma W, Hugenholtz J, Gregory JF III, Hanson AD. 2005a. A nudix enzyme removes pyrophosphate from dihydroneopterin triphosphate in the folate synthesis pathway of bacteria and plants. Journal of Biological Chemistry **280**, 5274-5280.

Klaus SMJ, Kunji ERS, Bozzo GG, Noiriel A, Diaz de la Garza R, Basset GJC, Ravanel S, Rébeillé F, Gregory JF III, Hanson AD. 2005b. Higher plant plastids and cyanobacteria have folate carriers related to those of Trypanosomatids. Journal of Biological Chemistry **280**, 38457-38463.

Kobayashi K, Forrest HS, el-Emary M. 1967. Isolation of 6-hydroxymethyllumazine (phosphodoxin?) and two other lumazine derivatives from spinach, and their synthesis. Archives of Biochemistry and Biophysics **121**:,220-223.

Kohashi M, Itadani T, Iwai K. 1980a. Purification and characterization of guanosine triphosphate cyclohydrolase I from *Serratia indica.* Agricultural and Biological Chemistry **44**, 271-278.
Kohashi M. 1980b. Isolation of six unconjugated pteridines from soybean (*Glycine max* L. Tsurunoko) seeds. Journal of Biochemistry(Tokyo) **91**, 1581-1586.

Kohashi M, Tomita K, Iwai K. 1980c. Analysis of unconjugated pterins in food resources and human urine. Agricultural and Biological Chemistry **44**, 2089-2094.

Kouzarides T. 2002. Histone methylation in transcriptional control. Current Opinion in Genetics and Development **12**, 198-209.

Kuhn DM, Lovenberg W. 1985. Tryptophan hydroxylase. **In**: *Folates and Pterins, Vol. 2.* pp. 115-154. Blackley, R.L. and Benkovic, S., Eds., J. Wiley (Interscience), New York.

Lewis SJ, Harbord RM, Harris R, Smith GD. 2006. Meta-analyses of observational and genetic association studies of folate intakes or levels and breast cancer risk. Journal of the National Cancer Institute **98**, 1607-1622.

Li Q, Curtis K, McCluskey J, Wahl T. 2002. Consumer attitudes toward genetically modified foods in Beijing, China. AgBioForum. **5**, 145-152.

Li ZW, Ren AG, Zhang L, Guo ZY, Li Z. 2006a. A population-based case-control study of risk factors for neural tube defects in four high-prevalence areas of Shanxi province, China. Paediatric and Perinatal Epidemiology **20**, 43-53.

Li ZW, Ren AG, Zhang L, Ye RW, Li S, Zheng JC, Hong SX, Wang TM, Li Z. 2006b. Extremely high prevalence of neural tube defects in a 4-county area in Shanxi Province, China. Birth Defects Research, Part A **76**, 237-240.

Li Z, Ren A, Zhang L, Liu J, Li Z. 2007. Periconceptional use of folic acid in Shanxi Province of northern China. Public Health Nutrition **10**, 471-476.

Lin M, Young CM. 2000. Biosynthesis of folates by *Streptococcus thermophilus* and *Lactobacillus delbrueckii* ssp. *bulgaricus*. Journal of Food and Drug Analysis **8**, 195-199.

Lin S, Rogiers S, Cossins EA. 1993. Gamma-Glutamyl hydrolase from pea cotyledons. Phytochemistry **32**, 1109-1117.

Lin XL, White RH. 1988. Structure of solfapterin (erythro-neopterin-3'-D-2-deoxy-2-aminoglucopyranoside) isolated from the thermophilic archaebacterium *Sulfolobus solfataricus*. Journal of Bacteriology **170**, 1396-1398.

Loizeau K, Gambonnet B, Zhang GF, Curien G, Jabrin S, Van Der Straeten D, Lambert WE, Rébeillé F, Ravanel S. 2007. Regulation of one-carbon metabolism in *Arabidopsis*: the N-terminal regulatory domain of cystathionine gamma-synthase is cleaved in response to folate starvation. Plant Physiology **145**, 491-503.

Loizeau K, De Brouwer V, Gambonnet B, Yu A, Renou JP, Van Der Straeten D, Lambert WE, Rébeillé F, Ravanel S. 2008. A genome-wide and metabolic analysis determined the adaptive response of *Arabidopsis* cells to folate depletion induced by methotrexate. Plant Physiology **148**, 2083-2095.

Lorence A, Chevone BI, Mendes P, Nessler CL. 2004. Myo-Inositol oxygenase offers a possible entry point into plant ascorbate biosynthesis. Plant Physiology **134**, 1200-1205.

Lowry OH, Bessey OA, Crawford EJ. 1949. Photolytic and enzymatic transformations of pteroylglutamic acid. Journal of Biological Chemistry **180**, 389-398.

Lucock M, Yates Z, Glanville T, Leeming R, Simpson N, Daskalakis I. 2003. A critical role for B-vitamin nutrition in human developmental and evolutionary biology. Nutrition Research **23**, 1463-1475.

Luo MZ, Orsi R, Patrucco E, Pancaldi S, Cella R. 1997. Multiple transcription start sites of the carrot dihydrofolate reductase thymidylate synthase gene, and sub-cellular localization of the bifunctional protein. Plant Molecular Biology **33**, 709-722.

Maqbool SB, Christou P. 1999. Multiple traits of agronomic importance in transgenic indica rice plants: analysis of transgene integration patters, expression levels and stability. Molecular Breeding **5**, 471-480.

Marean A, Graf A, Zhang Y, Niswander L. 2011. Folic acid supplementation can adversely affect murine neural tube closure and embryonic survival. Human Molecular Genetics **20**, 3678-3683.

Mathis JB, Brown GM. 1980. Dihydroneopterin aldolase from *Escherichia coli*. Methods in Enzymology **66**, 556-560.

McGuire JJ, Coward JK. 1984. Pteroylpolyglutamates: biosynthesis, degradation and function. **In**: *Folates and Pterins, Vol. 2.* pp. 135-190. Blackley, R.L. and Benkovic, S., Eds.,J. Wiley (Interscience), New York.

Melse-Boonstra A, Verhoef P, Konings EJM, van Dusseldorp M, Matser A, Hollman PCH, Meyboom S, Kok FJ, West CE. 2002. Influence of processing on Total, monoglutamate and polyglutamate folate contents of leeks, cauliflower, and green beans. Journal of Agriculture and Food Chemistry **50**, 3473-3478.

Mills JL, Von Kohorn I, Conley MR, Zeller JA, Cox C, Williamson RE, Dufour DR. 2003. Low vitamin B-12 concentrations in patients without anemia: the effect of folic acid of grain. American Journal of Clinical Nutrition **77**, 1474-1477.

Mitsuda H, Suzuki Y. 1971. Enzymatic conversion of 2-amino-4-hydroxy-6-formyl-7,8-dihydropteridine to 2-amino-4-hydroxy-6-hydroxymethyl-7,8-dihydropteridine by cell-free extracts of *Escherichia coli* B. Journal of Vitaminology **17**, 5-9.

Molloy A, Scott J. 2001. Folates and prevention of disease. Public Health Nutrition **4**, 601-609.

Moore L, Bradlee M, Singer M, Rothman K, Milunsky A. 2003. Folate intake and the risk of Neural Tube Defects: An estimation of Dose-Response. Epidemiology **14**, 200-205.

Morris MS, Jacques PF, Rosenberg IH, Selhub J. 2007. Folate and vitamin B-12 status in relation to anemia, macrocytosis, and cognitive impairment in older Americans in the age of folic acid fortification. American Journal of Clinical Nutrition **85**, 193-200.

Mouillon JM, Ravanel S, Douce R, Rébeillé F. 2002. Folate synthesis in higher plant mitochondria: coupling between the dihydropterin pyrophosphokinase and the dihydropteroate synthase activities. Biochemical Journal **363**, 313-319.

Nakase M, Hotta H, Adachi T, Aoki N, Nakamura R, Masumura T, Tanaka K, Matsuda T. 1996. Cloning of the rice seed α-globulin-encoding gene: sequence similarity of the 5'-flanking region to those of the genes encoding wheat high-molecular-weight glutenin and barley D hordein. Gene **170**, 223-226.

Naqvi S, Zhu C, Farre G, Ramessar K, Bassie L, Breitenbach J, Conesa DP, Ros G, Sandmann G, Capell T, Christou P. 2009. Transgenic multivitamin corn through biofortification of endosperm with three vitamins representing three distinct metabolic pathways. Proceedings of the National Academy of Sciences of the United States of America **106**, 7762-7767.

Nare B, Hardy LW, Beverley SM. 1997. The roles of pteridines reductase 1 and dihydrofoloate reductase-thymidylate synthase in pteridines metabolism in the protozoan parasite *Leishmania major*. Journal of Biological Chemistry **272**, 13883-13891.

Navarrete O, Van Daele J, Stove C, Lambert W, Van Der Straeten D, Storozhenko S. 2012. A folate independent role for cytosolic HPPK/DHPS upon stress in Arabidopsis thaliana. Phytochemistry **73**, 23-33.

Neuburger M, Rébeillé F, Jourdain A, Nakamura S, Douce R. 1996. Mitochondria are a major site for folate and thymidylate synthesis in plants. Journal of Biological Chemistry **271**, 9466-9472.

Nichols BP, Siebold AM, Doktor SZ. 1989. Para-aminobenzoate synthesis from chorismate occurs in two steps. Journal of Biological Chemistry **264**, 8597-8601.

Noiriel A, Naponelli V, Gregory JF III, Hanson AD. 2007a. Pterin and folate salvage. Plants and *Escherichia coli* lack capacity to reduce oxidized pterins. Plant Physiology **143**, 1101-1109.

Noiriel A, Naponelli V, Bozzo GG, Gregory JF III, Hanson AD. 2007b. Folate salvage in plants: pterin aldehyde reduction is mediated by multiple non-specific aldehyde reductases. Plant Journal **51**, 378-389.

Nunes ACS, Kalkmann DC, Aragão FJL. 2009. Folate biofortification of lettuce by expression of a codon optimized chicken GTP cyclohydrolase I gene. Transgenic Research **18**, 661-667.

Orsomando G, Diaz de la Garza R, Green BJ, Peng M, Rea PA, Ryan TJ, Gregory JF III, Hanson AD. 2005. Plant γ-glutamyl hydrolases and folate polyglutamates: characterization, Compartmentation and co-occurrence in vacuoles. Journal of Biological Chemistry **280**, 28877-28884.

Orsomando G, Bozzo GG, Diaz de la Garza R, Basset GJ, Quinlivan EP, Naponelli V, Rébeillé F, Ravanel S, Gregory JF III, Hanson AD. 2006. Evidence for folate-salvage reactions in plants. Plant Journal **46**, 426-435.

Ortiz PJ. 1970. Dihydrofolate and dihydropteroate synthesis by partially purified enzymes from wild-type and sulfonamide-resistant *Pneumococcus*. Biochemistry **9**, 355-361.

Ouellette M, Drummelsmith J, El Fadili A, Kundig C, Richard D, Roy G. 2002. Pterin transport and metabolism in *Leishmania* and related trypanosomatid parasites. International Journal of Parasitology **32**, 385-398.

Papakostas GI, Cassiello CF, Iovieno N. 2012. Folates and S-adenosylmethionine for major depressive disorder. Canadian Journal of Psychiatry **57**, 406-413.

Peach C, Velten J. 1991. Transgene expression variability (position effect) of CAT and GUS reporter genes driven by linked divergent T-DNA promoters. Plant Molecular Biology **17**, 49-60.

Piironen V, Edelmann M, Kariluoto S, Bed Z. 2008. Folate in wheat genotypes in the HEALTHGRAIN diversity screen. Journal of Agriculture and Food Chemistry **56**, 9726-9731.

Pitcher RJ, Edwards MC. 1995. *The effects of domestic cooking and preparation techniques on the nutritional composition of vegetables.* R&D Report No 18, Campden & Chorleywood Food Research Association.

Quinlivan EP, McPartlin J, McNulty H, Ward M, Strain JJ, Weir DG, Scott JM. 2002. Importance of both folic acid and vitamin B12 in reduction of risk of vascular disease. The Lancet **359**, 227-228.

Quinlivan EP, Roje S, Basset G, Shachar-Hill Y, Gregory JF III, Hanson AD. 2003. The folate precursos *p*-aminobenzoate is reversibly converted to its glucose ester in the plant cytosol. Journal of Biological Chemistry **278**, 20731-20737.

Rader JI, Schneeman BO. 2006. Prevalence of neural tube defects, folate status, and folate fortification of enriched cereal-grain products in the United States. Pediatrics **11**, 71394-1399.

Raichaudhuri A, Peng M, Naponelli V, Chen S, Sánchez-Fernández R, Gu H, Gregory JF III, Hanson AD, Rea PA. 2009. Plant vacuolar ABC transporters that translocate folates and antifolates *in vitro* and contribute to antifolate tolerance *in vivo*. Journal of Biological Chemistry **284**, 8449-8460.

Ray JG, Cole DEC, Boss SC. 2000. An Ontario-wide study of vitamin B12, serum folate, and red cell folate levels in relation to plasma homocysteine: Is a preventable public health issue on the rise? Clinical Biochemistry **33**, 337-343.

Ray JG, Vermeulen MJ, Langman LJ, Boss SC, Cole DEC. 2003. Persistence of vitamin B12 insufficiency among elderly women after folic acid food fortification. Clinical Biochemistry **36**, 387-391.

Ravanel S, Cherest H, Jabrin S, Grunwald D, Surdin-Kerjan Y, Douce R, Rébeillé F. 2001. Tetrahydrofolate biosynthesis in plants: Molecular and functional characterization of dihydrofolate synthetase and three isoforms of folylpolyglutamate synthetase in *Arabidopsis thaliana*. Proceedings of the National Academy of Sciences of the United States of America **98**, 15360-15365.

Ravanel S, Block MA, Rippert P, Jabrin S, Curien G, Rébeillé F, Douce R. 2004. Methionine metabolism in plants: chloroplasts are autonomous for *de novo* methionine synthesis and can import S-adenosylmethionine from the cytosol. Journal of Biological Chemistry **279**, 22548-22557.

Ravanel S, Douce R, Rébeillé F. 2011. Metabolism of folates in plants. Advances in Botanical Research **59**, 67-106.

Razin A, Riggs AD. 1980. DNA methylation and gene function. Science **210**, 604-610.

Rébeillé F, Neuburger M, Douce R. 1994. Interactions between glycine decarboxylase, serine hydroxymethyltransferase and tetrahydrofolate polyglutamates in pea leaf mitochondria. Biochemical Journal **302**, 223-228.

Rébeillé F, Macherel D, Mouillon JM, Garin J, Douce R. 1997. Folate biosynthesis in higher plants: purification and molecular cloning of a bifunctional 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase/7,8-dihydropteroate synthase localized in mitochondria. EMBO Journal **16**, 947-957.

Rébeillé F, Ravanel S, Jabrin S, Douce R, Storozhenko S, Van Der Straeten D. 2006. Folates in plants: biosynthesis, distribution and enhancement. Physiologia Plantarum **126**, 330-342.

Reed LS, Archer MC. 1980. Oxidation of tetrahydrofolic acid by air. Journal of Agriculture and Food Chemistry **28**, 801-805.

Ren A, Zhang A, Li Z, Hao L, Tian Y, Li Z. 2006. Awareness and use of folic acid, and blood folate concentrations among pregnant women in northern China - An area with a high prevalence of neural tube defects. Reproductive Toxicology **22**, 431-436.

Reynolds EH. 2006. Vitamin B12, folic acid, and the nervous system. Lancet Neurology **5**, 949-960.

Roje S, Wang H, McNeil SD, Raymond RK, Appling DR, Shachar-Hill Y, Bohnert HJ, Hanson AD. 1999. Isolation, characterization, and functional expression of cDNAs encoding NADH-dependent methylenetetrahydrofolate reductase from higher plants. Journal of Biological Chemistry **274**, 36089-36096.

Roje S, Janave MT, Ziemak MJ, Hanson AD. 2002. Cloning and characterization of mitochondrial 5-formyltetrahydrofolate cycloligase from higher plants. Journal of Biological Chemistry **277**, 42748-42754.

Roos AJ, Cossins EA. 1971. Pteroylglutamate derivatives in *Pisum sativum* L. Biosynthesis of cotyledonary tetrahydropteroylglutamates during germination. Biochemical Journal **125**, 17-26.

Ross J, Green J, Baugh CM, MacKenzie RE, Matthews RG. 1984. Studies on the polyglutamate specificity of methylenetetrahydrofolate dehydrogenase from pig liver. Biochemistry **23**, 1796-1801.

Sahr T, Ravanel S, Basset G, Nichols BP, Hanson AD, Rébeillé F. 2006. Folate synthesis in plants: purification, kinetic properties, and inhibition of aminodeoxychorismate synthase. Biochemical Journal **396**, 157-162.

Sanjoaquin MA, Allen N, Couto E, Roddam AW, Key TJ. 2005. Folate intake and colorectal cancer risk: a meta-analytical approach. International Journal of Cancer **113**, 825-828.

Santos F, Wegkamp A, de Vos MD, Smid EJ, Hugenholtz J. 2008. High-level folate production in fermented foods by the B_{12} producer *Lactobacillus reuteri* JCM1112. Applied and Environmental Microbiology **74**, 3291-3294.

Scholl TO, Johnson WG. 2000. Folic acid: influence on the outcome of pregnancy. American Journal of Clinical Nutrition **71**, 1295S-1303S.

Scott JM. 1984. Catabolism of folates. **In**: *Folates and Pterins, Vol. 1.* pp. 307-327. Blackley, R.L. and Benkovic, S., Eds., J. Wiley (Interscience), New York.

Scott JM, Weir DG. 1986. Excretion of folates, pterins and their metabolic products. **In**: *Folates and Pterins, Vol. 3.Nutritional, Pharmacological and Physiological Aspects.* pp. 297-324. Blackley, R., Eds., J. Wiley and Sons Inc., New York.

Scott JM, Weir DG. 1996. Homocysteine and cardiovascular disease. The Quarterly Journal of Medecine **89**, 561-563.

Scott J, Rébeillé F, Fletcher J. 2000. Folic acid and folate: the feasibility for nutritional enhancement in plant foods. Journal of the Science of Food and Agriculture **80**, 795-824.

Seshadri S, Beiser A, Selhub J, Jacques PF, Rosenberg IH, D'Agostino RB, Wilson PWF, Wolf PA. 2002. Plasma homocysteine as a risk factor for dementia and Alzheimer's disease. The New England Journal of Medecine **346**, 476-483.

Shane B. 1989. Folylpolyglutamate synthesis and role in the regulation of one-carbon metabolism. Vitamins and Hormones **45**, 263-335.

Shanxi Province Statistical Bureau. 2007. *Shanxi Statistical Yearbook 2007.* Beijing, Shanxi Statistics Press.

Sherwood RF, Melton RG, Alwan SM, Hughes P. 1985. Purification and properties of carboxypeptidase G2 from *Pseudomonas* sp. Strain RS-16. European Journal of Biochemistry **148**, 447-453.

Shin YS, Kim ES, Watson JE, Stokstad EL. 1975. Studies of folic acid compounds in nature. IV. Folic acid compounds in soybeans and cow milk. Can. Journal of Biochemistry**53**, 338-343.

Shiota T. 1959. Enzymic synthesis of folic acid-like compounds by cell-free extracts of *Lactobacillus arabinosus*. Archives of Biochemistry and Biophysics **80**, 155-161.

Shiota T, Baugh CM, Jackson R, Dollard R. 1969. The enzymatic synthesis of hydroxymethyldihydropteridine pyrophosphate and dihydrofolate. Biochemistry **8**, 5022-5028.

Siehl D. 1998. The biosynthesis of tryptophan, tyrosine and phenylalanine from chorismate. **In**: *Plant Amino Acids.* pp. 171-204. Singh, B.K., Ed., Marcel Dekker Inc., New York.

Singh S, Sedgh G, Hussain R. 2010. Unintended pregnancy: worldwide levels, trends, and outcomes. Studies in Family Planning **41**, 241-250.

Smith AD, Refsum H. 2008. Folic acid supplements are good (not bad) for rheumatoid arthritis patients treated with low-dose methotrexate – Reply. American Journal of Clinical Nutrition **88**, 480-480.

Sohta Y, Ohta T, Masada M. 1997. Purification and some properties of GTP cyclohydrolase I from spinach leaves. Bioscience, Biotechnology and Biochemistry **61**, 1081-1085.

Song J, Medline A, Mason JB, Gallinger S, Kim YI. 2000a. Effects of dietary folate on intestinal tumorigenesis in the apcMin mouse. Cancer Research **60**, 5434-5440.

Song J, Sohn KJ, Medline A, Ash C, Gallinger S, Kim YI. 2000b. Chemopreventive effects of dietary folate on intestinal polyps in Apc+/-Msh2-/- mice. Cancer Research **60**, 3191-3199.

Srivastava AC, Ramos-Parra PA, Bedair M, Robledo-Hernandez AL, Tang YH, Sumner LW, de la Garza RID, Blancaflor EB. 2011. The folylpolyglutamate plastidial isoform is required for post-embryonic root development in Arabidopsis. Plant Physiology **155**, 1237-1251.

Stakhov LF, Ladygin VG, Stakhova LN. 2002. Effect of ultraviolet C irradiation on folate and free amino acid contents in leaves of *Pisum sativum*. Biofizika **47**, 878-885.

Stanger O. 2004. The potential role of homocysteine in percutaneous coronary interventions (PCI): Review of current evidence and plausibility of action. Cellular Molecular Biology **50**, 953-988.

Stokes ME, Chattopadhyay A, Wilkins O, Nambara E, Campbell MM. 2013. Interplay between sucrose and folate modulates auxin signaling in Arabidopsis. Plant Physiology **162**, 1552-1565.

Storozhenko S, Ravanel S, Zhang GF, Rébeillé F, Lambert W, Van Der Straeten D. 2005. Folate enhancement in staple crops by metabolic engineering. Trends in Food science and Technology **16**, 271-281.

Storozhenko S, Navarrete O, Ravanel S, De Brouwer V, Chaerle P, Zhang G-F, Bastien O, Lambert W, Rébeillé F, Van Der Straeten D. 2007a. Cytosolic hydroxymethyldihydropterin pyrophosphokinase/dihydropteroate synthase from *Arabidopsis thaliana*. Journal of Biological Chemistry **282**, 10749-10761.

Storozhenko S, De Brouwer V, Volckaert M, Navarrete O, Blancquaert D, Zhang G-F, Lambert W, Van Der Straeten D. 2007b. Folate fortification of rice by metabolic engineering. Nature Biotechnology **25**, 1277-1279.

Stover P, Schirch V. 1993. The metabolic role of leucovorin. Trends in Biochemical Sciences **18**, 102-106.

Suh JR, Oppenheim EW, Girgis S, Stover PJ. 2000. Purification and properties of a folate-catabolizing enzyme. Journal of Biological Chemistry **275**, 35646-35655.

Suh JR, Herbig AK, Stover PJ. 2001. New perspectives on folate catabolism. Annual Review of Nutrition **21**, 255-282.

Suzuki Y, Brown GM. 1974. The biosynthesis of folic acid XII. Purification and properties of dihydroneopterin triphosphate pyrophosphohydrolase. Journal of Biological Chemistry **249**, 2405-2410.
Swedberg G, Castensson S, Sköld O. 1979. Characterization of mutationally altered dihydropteroate synthase and its ability to form a sulfonamide-containing dihydrofolate analog. Journal of Bacteriology **137**, 129-136.

Sybesma W, Starrenburg M, Kleerebezem M, Mierau I,de Vos WM, Hugenholtz J. 2003a. Increased production of folate by metabolic engineering of *Lactococcus lactis*. Applied and Environmental Microbiology **69**, 3069-3076.

Sybesma W, van den Born E, Starrenburg M, Mierau I, Kleerebezem M, de Vos WM, Hugenholtz J. 2003b. Controlled modulation of folate polyglutamyl tail length by metabolic engineering of *Lactococcus lactis*. Applied and Environmental Microbiology **69**, 7101-7107.

Takaiwa F, Oono K, Wing D, Kato A. 1991. Sequence of three members and expression of a new major subfamily of glutelin genes from rice. Plant Molecular Biology **17**, 875-885.

Takimoto H, Hayashi F, Kusama K, Kato N, Yoshiike N, Toba M, Ishibashi T, Miyasaka N, Kubota T. 2011. Elevated maternal serum folate in the third trimester and reduced fetal growth: a longitudinal study. Journal of Nutritional Science and Vitaminology **57**, 130-137.

Tate PH, Bird AP. 1993. Effects of DNA methylation on DNA-binding proteins and gene expression. Current Opinion in Genetics and Development **3**, 226-231.

Tewari YB, Jensen PY, Kishore N, Mayhew MP, Parsons JF, Eisenstein E, Goldberg RN. 2002. Thermodynamics of reactions catalyzed by PABA synthase. Biophysical Chemistry **96**, 33-51.

Toth I, Lazar G, Goodman HM. 1987. Purification and immunochemical characterization of a dihydrofolate-reductase-thymidylate synthase enzyme complex from wild-carrot cells. EMBO Journal **6**, 1853-1858.

Triglia T, Cowman AF. 1994. Primary structure and expression of the dihydropteroate synthetase gene of *Plasmodium falciparum*. Proceedings of the National Academy of Sciences of the United States of America **91**, 7149-7153.

USDA U.S. Department of Agriculture, Agricultural Research Service. 2012. USDA National Nutrient Database for Standard Reference, Release 25. Nutrient Data Laboratory Home Page, http://www.ars.usda.gov/ba/bhnrc/ndl.

Van Wilder V, De Brouwer V, Loizeau K, Gambonnet B, Albrieux C, Van Der Straeten D, Lambert WE, Douce R, Block MA, Rébeillé F, Ravanel S. 2008. C1 metabolism and chlorophyll synthesis: the Mg-protoporphyrin IX methyltransferase activity is dependent on the folate status. New Phytologist **182**, 137-145.

Vauclare P, Diallo N, Bourguignon J, Macherel D, Douce R. 1996. Regulation of the expression of the glycine decarboxylase complex during pea leaf development. Plant Physiology **112**, 1523-1530.

Wald NJ, Watt HC, Law MR, Weir DG, McPartlin J, Scott JM. 1998. Homocysteine and ischaemic heart disease: results of a prospective study with implications on prevention. Archives of Internal Medicine **158**, 862-867.

Wald DS, Wald NJ, Morris JK, Law M. 2006. Folic acid, homocysteine, and cardiovascular disease: judging causality in the face of inconclusive trial evidence. British Medical Journal **333**, 1114-1117.

Waller JC, Akhtar TA, Lara-Núñez A, Gregory JF III, McQuinn RP, Giovannoni JJ, Hanson AD. 2010. Developmental and feedforward control of the expression of folate biosynthesis genes in tomato fruit. Molecular Plant **3**, 66-77.

Wang J, Leblanc E, Chang C-F, Papadopoulou B, Bray T, Whiteley JM, Lin S-X, Ouellette M. 1997. Pterin and folate reduction by the *Leishmania tarentolae* H locus short-chain dehydrogenase/reductase PTR1. Archives of Biochemistry and Biophysics **342**, 197-202.

Ward M, McNulty H, McPartlin J, Strain JJ, Weir DG, Scott JM. 1997. Plasma homocysteine, a risk factor for cardiovascular disease, is lowered by physiological doses of folic acid. The Quarterly Journal of Medicine **90**, 519-524.

Weaver LM, Herrmann KM. 1997. Dynamics of the shikimate pathway in plants. Trends in Plants Sciences **2**, 346-351.

Wegkamp A, Starrenburg M, de Vos WM, Hugenholtz J, Sybesma W. 2004. Transformation of folate-consuming *Lactobacillus gasseri* into a folate producer. Applied and Environmental Microbiology **70**, 3146-3148.

Wegkamp A, van Oorschot W, de Vos WM, Smid EJ. 2007. Characterization of the role of *para*-aminobenzoic acid biosynthesis in folate production by *Lactococcus lactis*. Applied and Environmental Microbiology **73**, 2673-2681.

Whiteley JM, Drais J, Kirchner J, Huennekens FM. 1968. Synthesis of 2-amino-4-hydroxy-6-formyl-7,8-dihydropteridine and its identification as a degradation product of dihydrofolate. Archives of Biochemistry and Biophysics **126**, 955-957.

Witter K, Werner T, Blusch JH, Schneider EM, Reiss O, Ziegler I, Rodl W, Bacher A, Gutlich M. 1996. Cloning, sequencing and functional studies of the gene encoding human GTP cyclohydrolase I. Gene **171**, 285-290.

Ye Q-Z, Liu Y, Walsh CT. 1990. *p*-Aminobenzoate synthesis in Escherichia coli: purification and characterization of PabB as aminodeoxychorismate synthase and enzyme X as aminodeoxychorismate lyase. Proceedings of the National Academy of Sciences of the United States of America **87**, 9391-9395.

Yoneyama T, Hatakeyama K. 1998. Decameric GTP cyclohydrolase I forms complexes with two pentameric GTPCH feedback regulatory proteins in the presence of phenylalanine or of a combination of BH 4 and GTP. Journal of Biological Chemistry **273**, 20102-20108.

Zhang B, Zhang T, Lin L, Wang F, Xin R, Gu X, He Y, Yu D, Li P, Zhang Q, Zhao J, Qin Y, Yang X, Chen G, Liu J, Song X, Zheng X. 2008. Correlation Between Birth Defects and Dietary Nutrition Status in a High Incidence Area of China. Biomedical and Environmental Science **21**, 37-44.

Zhou J, Kang S, Wong PWK, Fournier B, Rozen R. 1990. Purification and characterization of methylenetetrahydrofolate reductase from human cadaver liver. Biochemical Medicine and Metabolic Biology **43**, 234-242.

Zhou HR, Zhang FF, Ma ZY, Huang HW, Jiang L, Cai T, Zhu JK, Zhang C, He XJ. 2013. Folate polyglutamylation is involved in chromatin silencing by maintaining global DNA methylation and histone H3K9 dimethylation in Arabidopsis. Plant Cell **25**, 2545-2559.

Zhu T, Pan Z, Domagalski N, Koepsel R, Ataai MM, Domach MM. 2005. Engineering of *Bacillus subtilis* for enhanced total synthesis of folic acid. Applied and Environmental Microbiology **71**, 7122-7129.

Zrenner R, Stitt M, Sonnewald U, Boldt R. 2006. Pyrimidine and purine biosynthesis and degradation in plants. Annual Reviews in Plant Biology **57**, 805-836.

SCOPE

Folates are essential vitamins (B9), which play an important role in C1 metabolism in both eukaryotic and prokaryotic organisms. They are key molecules in the methyl cycle and the *de novo* biosynthesis of e.g. nucleic acids. Folate deficiency in humans can result in the onset of several diseases and disorders. Only plants and micro-organisms are able to synthesize folates *de novo* and folate content in most staple crops is very low, especially in rice. Since rice is the major food source for more than half of world's population, consequences of folate deficiency are very common within these population groups. A couple of years ago, transgenic rice lines (GA-lines) with a 100-fold higher folate content than wild type controls were created at Ghent University. This biofortification resulted from the overexpression of two *Arabidopsis* folate biosynthesis genes: GTP cyclohydrolase I (GTPCHI) and amino-deoxychorismate synthase (ADCS).

In the first part of this work, we investigated the effect of folate enhancement on general rice seed metabolism. Therefore, we sampled rice seeds from early on in development (day 8 after anthesis) up until the end of seed maturation (28 days after anthesis), with an interval of 4 days. We investigated the accumulation of folate and its precursors p-ABA and pterin in WT, A, G and two GA lines (GA26.5 and GA9.15) and measured the expression of the transgenes in those transgenic lines during rice seed development. Subsequently, we chose 3 time points for a microarray hybridization experiment, to investigate the effect of folate accumulation on developing GA vs. WT rice seeds. This way, we revealed a number of genes which were affected in expression by folate enhancement. Also, we investigated the effect of folate accumulation on the expression of endogenous folate biosynthesis genes. In the second part of this work, we tried to enhance folate levels in *Arabidopsis* plants and potato tubers by metabolic engineering. Therefore, the same two-gene strategy, successful in rice and tomato folate biofortification, was used for both plant species. Several constructs were made for *Arabidopsis thaliana* and potato transformation, in order to enhance p-ABA and pterin levels and a p-ABA feeding experiment was conducted on *Arabidopsis* seedlings, overexpressing GTPCHI, the first enzyme in the pterin branch of the folate biosynthesis pathway. In the last part of this work, we tried to enhance folate stability in biofortified rice seeds. To do so, we decided to follow two strategies: increasing folate polyglutamylation, which promotes cellular retention and binding to folate dependent enzymes, and enhancement of folate stability through complexation with folate binding proteins (FBP). A total of 11 rice transformation constructs were made, on which the two genes (GTPCHI and ADCS), necessary for folate enhancement, were combined with one or both strategies. To enhance folate polyglutamylation, two *Arabidopsis thaliana* isoforms of folylpolyglutamate synthetase (FPGS), i.e. mitochondrial and cytosolic FPGS, were cloned for overexpression in rice endosperm. To bind the accumulating folates in biofortified rice seeds, a synthetic, soluble FBP, based on bovine FBP, was created and overexpressed in rice seeds on itself or as a fusion with rice glutelin B4 or *Arabidopsis thaliana* β carbonic anhydrase 2, two proteins which are abundant in rice seeds and plant cell cytosol respectively. Next to this, we investigated the effect of temperature during long storage on folate stability in WT, GA lines and the newly created transgenics.

Chapter 1

Rice folate enhancement through metabolic engineering has an impact on rice seed metabolism, but does not affect the expression of the endogenous folate biosynthesis genes

All the experiments were performed by Dieter Blancquaert, folate, p-ABA and pterin analyses were performed by Jeroen Van Daele (Faculty of Pharmaceutical Sciences, Laboratory of Toxicology).

Rice folate enhancement through metabolic engineering has an impact on rice seed metabolism, but does not affect the expression of the endogenous folate biosynthesis genes

Dieter Blancquaert[a], Jeroen Van Daele[b], Sergei Storozhenko[a], Christophe Stove[b], Willy Lambert[b], Dominique Van Der Straeten[a]

[a] Laboratory of Functional Plant Biology, Department of Physiology, Ghent University, K.L. Ledeganckstraat 35, 9000 Gent, Belgium

[b] Laboratory of Toxicology, Ghent University, Harelbekestraat 72, 9000 Gent, Belgium

ABSTRACT

Folates are key-players in one-carbon metabolism in all organisms. However, only micro-organisms and plants are able to synthesize folates *de novo* and humans rely entirely on their diet as a sole folate source. As a consequence, folate deficiency is a global problem. Although different strategies are currently implemented to fight folate deficiency, up until now, all of them have their own drawbacks. As an alternative and complementary means to those classical strategies, folate biofortification of rice by metabolic engineering was successfully achieved a couple of years ago. To gain more insight into folate biosynthesis regulation and the effect of folate enhancement on general rice seed metabolism, a transcriptomic study was conducted in developing transgenic rice seeds, overexpressing 2 genes of the folate biosynthetic pathway. Upon folate enhancement, the expression of 235 genes was significantly altered. Here, we show that rice folate biofortification has an important effect on folate dependent, seed developmental and plant stress response/defense processes, but does not affect the expression of the endogenous folate biosynthesis genes.

Adapted from:

Blancquaert D, Van Daele J, Storozhenko S, Stove C, Lambert W, Van Der Straeten D. 2013. Rice folate enhancement through metabolic engineering has an impact on rice seed metabolism, but does not affect the expression of the endogenous folate biosynthesis genes. Plant Molecular Biology **83**, 329-349.

Introduction

Folates comprise a group of water-soluble B-vitamins (B9) which are indispensable cofactors in one-carbon (C1) metabolism, where they act as both one-carbon donors and acceptors (Blancquaert et al., 2010). They play an important role in DNA biosynthesis, through their involvement in purine and thymidylate formation (Smith et al., 1980), in serine, glycine (Besson et al., 1993) and pantothenate (vitamin B5) biosynthesis (Asensi-Fabado and Munné-Bosch 2010), in the degradation of histidine (Cook 2001) and in methionine metabolism (Krebs et al., 1976). In addition, they are involved in the methyl cycle, by providing S-adenosylmethionine to methyltransferases, which, in turn, can methylate a set of substrates, such as hormones, DNA, lipids and proteins (Scott et al., 2000). Furthermore, the methyl cycle is important in the catabolism of excess dietary methionine (Lucock et al., 1996) and the regulation of gene expression (Tate and Bird 1993). Additionally, folates play a role in photorespiration, N metabolism (Jiang et al., 2013) and in the biosynthesis of chlorophyll, betaine, alkaloids and lignin in plants (Hanson and Roje 2001). Folates are composed of three parts: para-aminobenzoate (p-ABA), a pterin moiety, and a glutamate tail of one to fourteen γ-linked glutamates. Folates can differ in their oxidation state, the C1 unit attached to the molecule and the extent of polyglutamylation. Thus, a pool of different folate forms is present in cells, each having specific functions in C1 metabolism. Since folates can only be synthesized *de novo* by plants and micro-organisms, humans are entirely dependent on their diet as the sole source of folates. Green leafy vegetables, leguminous vegetables, certain fruits (e.g. strawberries and oranges), and fermented products are rich folate sources. However, most staple crops, such as potato, banana, maize, rice and other cereals have a poor folate content (USDA National Nutrient Database for Standard References; http://ndb.nal.usda.gov/). Hence, folate deficiency is an important problem, affecting both developing and developed regions all over the world. A low folate status can cause the onset of diseases and disorders, such as neural tube defects (e.g. anencephaly and spina bifida) (Geisel 2003) and megaloblastic anaemia (Li et al., 2003). Moreover, some neurodegenerative disorders, such as Alzheimer disease (Seshadri et al., 2002), major depressive disorder (MDD) (Papakostas et al., 2012) and a higher risk of cardiovascular (Scott and Weir 1996) and coronary diseases (Stanger 2004), stroke (Endres et al., 2005) and a wide range of cancers (Choi and Friso 2005) have been associated with a low folate status. Up until now, several strategies have been applied to fight folate deficiency: supplementation by folic acid pills, folic acid food fortification and dietary changes (for a review, see Blancquaert et al., 2013). However, each of these has its own drawbacks and limitations and these strategies are difficult to implement to the poorer populations in the world. Therefore, folate biofortification, the enhancement of natural folate content in staple crops through plant breeding or metabolic engineering, has been suggested to offer a sustainable alternative to the above-mentioned strategies (Bekaert et al., 2008). In this respect, *japonica* Nipponbare rice has been successfully biofortified with folates by overexpressing *Arabidopsis thaliana* cDNA encoding GTP cyclohydrolase I (GTPCHI, G) and ADC synthase (ADCS, A), the first enzymes in pterin and p-ABA biosynthesis respectively, both under the control of a rice endosperm specific promoter (resulting in transgenic GA rice lines) (Storozhenko et al., 2007). By then, it was decided to overexpress *Arabidopsis* cDNA instead of rice genes, because rice folate biosynthesis genes were not completely characterized in contrast to their *Arabidopsis* counterparts. A hundredfold enhancement of natural folate levels, as compared to wild type (WT) rice seeds, has been achieved in these lines, with a total folate content up to 1723 µg per 100 g fresh weight (FW) (Storozhenko et al., 2007). Since the recommended daily intake (RDI) of folates is 400 µg for an adult (and 600µg for a pregnant woman) (National Institutes of

Health - Office of Dietary Supplements; http://ods.od.nih.gov/factsheets/Folate-HealthProfessional/), the folate content of 100 g of this transgenic rice is four-fold the RDI for an adult person. Taking into account that folate bioavailability (Bailey 2004) and cooking (Storozhenko et al., 2007) further quarters the amount of folates that can be effectively used by the human body, this GA rice can become an ideal vehicle in the battle against folate deficiency. Since C1 metabolism, in which folates are of paramount importance, plays an important role in many aspects of plant physiology, it is not excluded that enhancement of folate levels could have a major impact on other physiological processes in a biofortified crop. Therefore, it is necessary to investigate the effects of increased folate content on general metabolism in the folate biofortified rice and uncover gene clusters and their corresponding cellular processes which may be affected by high folate levels. Secondly, since little is known about the regulation of the folate biosynthesis pathway in plants, it is interesting to investigate the effect of *Arabidopsis GTPCHI* and *ADC synthase* overexpression and the resulting accumulation of pterins, p-ABA and folates on the different endogenous rice folate biosynthesis genes. In this study, an extensive transcriptomic profiling by microarray hybridization was carried out on developing rice seeds to investigate the effect of folate enhancement on general rice seed metabolism.

Material and methods

Plant material and sampling

Five *Oryza sativa* ssp. *japonica* var. Nipponbare lines were grown in soil in short day light conditions (8 hours of light (420 µmoles/m^2.s); 16 hours of darkness) at a constant temperature and humidity of 28°C and 80% respectively. These five lines comprised one wild type (WT) line and four transgenic rice lines; a G-line (2G 17.1.5.5; overexpressing *Arabidopsis thaliana* cDNA encoding GTP cyclohydrolase 1, under the control of the endosperm-specific globulin promoter); an A-line (2A 11.2.7.5; overexpressing *Arabidopsis thaliana* cDNA encoding ADC synthase, under the control of the endosperm-specific glutelin B1 promoter) and two GA lines, GA26.5 (with intermediate folate levels of 12.02 ± 4.53 nmol/g fresh weight (FW)) and GA9.15 (with high folate content of 38.3 ± 0.16 nmol/g FW) (two different promoters were used to avoid promoter interference in the GA lines). All four transgenic lines resulted from a transformation with a single T-DNA construct, combining the abovementioned transgenes *GTP cyclohydrolase 1* and *ADC synthase*, under their endosperm-specific promoters (Storozhenko et al., 2007) and were confirmed to contain a single copy of the T-DNA, in a homozygous state. Plants were grown simultaneously in two batches, in randomized order. Upon anthesis, seeds were monitored and labeled. Sampling of the seeds occurred at 8, 12, 16, 20, 24 and 28 days post-anthesis (DPA). Seeds from each time point and from each line were manually dehusked and the resulting samples were kept at -80°C until further analysis. Approximately 45 seeds were collected per time point in each batch.

Folate, p-ABA and pterin analysis

All chemicals, materials, stock and working solutions used and prepared for folate, p-ABA and pterin analyses were as previously described (De Brouwer et al., 2008; Navarrete et al., 2012).

Sample preparation

During sample preparation for folate and pterin analyses, all manipulations were carried out under subdued light.

For folate determination, 1.5 ml of extraction buffer (containing the internal standards, see De Brouwer et al., 2008 for buffer composition) was added to ± 200 mg of rice seeds and the capped Eppendorf® tube was placed at 100°C for 10 minutes before homogenization in a retsch mill. A tri-enzyme treatment with 30 µl α-amylase (23.5 units/µl) (20 minutes), 150 µl protease (≥3.5 units/mg) (1 hour at 37 °C) and 100 µl conjugase (2 hours at 37°C) was used to degrade the starch matrix, to release protein-bound folates and to deconjugate polyglutamylated folates. To stop protease and conjugase activity, additional heat treatments were carried out, again followed by cooling on ice. The resulting solution was ultra-filtrated at 12851 x *g* for 15 min.

For pterin analysis, the procedure was equal to that of folate determination, with the exception of an additional 1 hour incubation at 37°C with 20 µl α-amylase and protease, followed by a heat treatment and ultra-filtration.

For p-ABA analysis, the same amount of rice seeds were ground after boiling in only 200 µl of deionized water. Afterwards, the samples were transferred to 15-ml Falcon® tubes, meanwhile diluting these with a 11.8 ml volume of methanol containing 3-amino-4-methylbenzoic acid. After extraction, the tubes were centrifuged at 2655 x g for 15 min. The supernatant was collected and 5 ml of the methanolic layer was dried completely under a nitrogen gas flow at 45°C. A 1 ml aliquot of water was added to the residue, followed by sonication for 5 min. To a volume of 0.4 ml of this solvated residue, 50 µl of 2 M HCl was added. After capping, the tube was incubated at 80°C for 2 hours and after cooling of the solution, 50 µl of 2 M NaOH was added for neutralization. Finally, the sample solutions were ultra-filtrated.

UPLC™ conditions

The chromatography for folate, p-ABA and pterin analyses was performed as previously described (De Brouwer et al., 2010; Navarrete et al., 2012), with minor modifications: for pterin analysis (column: 30°C), the mobile phase consisted of eluent A (0.1% acetic acid in water) and eluent B (0.1% acetic acid in acetonitrile) and was pumped at a flow rate of 0.45 ml/min. The starting condition (100% A) was decreased linearly to 95% in 2.5 minutes, followed by a nonlinear decrease to 0% in 1.5 minutes, where it was kept for 1 minute. Subsequently, the mobile phase was adjusted to its initial composition, again in a nonlinear way, and held for 2 minutes for re-equilibration, resulting in a total time of 7 minutes.

Mass spectrometric instrumentation and settings

The detection for the different analyses was performed by ESI utilizing heated auxiliary gas (Turbospray), in the *scheduled* MRM™ mode, using an Applied Biosystems API 4000 tandem mass spectrometer (Ontario, Canada). For folate analysis, the source conditions were optimized and were set as follows: source temperature at 500°C, ionspray voltage at 3.5 kV, Q1 and Q3 at unit mass resolution. The interface heater was on; gas 1, gas 2, collision activated dissociation (CAD) gas and curtain gas were nitrogen, with pressure settings at 70, 90, 8 and 25, respectively. For pterin analysis

the settings were as follows: source temperature at 500°C, ionspray voltage at -4.5 kV, Q1 and Q3 at unit mass resolution. The interface heater was on; gas 1, gas 2, curtain gas and CAD gas were nitrogen, with settings at 80, 90, 26, and 8, respectively. For p-ABA analysis the settings were as follows: source temperature was set at 550°C, ionspray voltage at 5.5 kV, Q1 and Q3 at unit mass resolution. The interface heater was on; gas 1, gas 2, curtain gas and CAD gas were nitrogen, with settings at 75, 20, 30, and 6, respectively.

The compound-specific parameters for folate, pterin and p-ABA analyses are listed in Supplemental Tables S1.1, S1.2 and S1.3 respectively.

RNA extraction, cDNA synthesis and real-time quantitative PCR

Total RNA was extracted using Trizol™ reagent (Invitrogen) according to the manufacturer's instructions, with minor modifications (after tissue homogenization, samples were incubated at room temperature in Trizol™ for one hour under continuous shaking) , and subsequently purified (RNeasy Plant Mini Kit, Qiagen) and treated with DNase (RNase-Free DNase Set, Qiagen), according to the manufacturer's instructions, with minor modifications (the RNA pellet obtained after Trizol™ purification was dissolved in buffer RLC prior to column loading). One ug of RNA was reverse-transcribed to cDNA by using the Verso cDNA Kit (Thermo Scientific) with oligo-dT primers. The resulting cDNA was diluted 5 times with deionized water. Real-time quantitative PCR was performed using the iCycler (Bio-Rad) as a thermal cycler and IQ™ 5 (Bio-Rad) as the optical module for multicolor real-time quantitative PCR detection. MAXIMA™ SYBR Green qPCR Master Mix (Fermentas) was used in a 25 ul reaction mixture of 5 ul of diluted cDNA and 800 nM of each primer (Table 1.1). All samples included two technical replicates. Each real-time quantitative PCR run comprised 40 cycles (55°C annealing, 72°C extension and 95°C denaturation temperature). Data analysis was performed using qBASE, a software program which calculates the relative data based on the $2^{\Delta\Delta Ct}$ method (Hellemans et al., 2007; Livak et al., 2001).

Microarray hybridization and data analysis

After the appropriate quality controls (NASC's International Affymetrix Service, Nottingham), 1 ug of RNA of WT, GA26.5 and GA9.5 of 8, 12 and 20 DPA of both batches was used for transcriptomic analysis by microarrays. Upon cDNA labeling, the samples were subjected to GeneChip® Rice Genome Array microarray hybridization (NASC's International Affymetrix Service, Nottingham). This array contains probes to query 48,564 and 1,260 transcripts of the *japonica* and *indica* subspecies respectively. Microarray quality assessment and data analysis were performed with Robin, a user-friendly R-based software program (Lohse et al., 2010). Affymetrix CEL files were imported into Robin and the Robust Multichip Average (RMA) algorithm was used to create an expression matrix from the Affymetrix data. This algorithm performs subsequently a background correction, a quantile normalization and a log2 transformation of the raw intensity values (Irizarry et al., 2003). Within each time point, the normalized expression data of both GA lines (batch 1 + batch 2) were grouped and compared to those of WT (batch 1 + batch 2). Differentially expressed genes in the GA lines were identified, using Benjamini and Hochberg (1995) for P-value correction (false discovery rate), nested F as a multiple testing strategy, a minimal log fold change of 1 (compared to the control) and 0.05 as a corrected P-value cut-off.

Table 1.1 List of primers used for real-time quantitative PCR.

Function	Locus	Forward primer	Reverse primer
GTPCHI Arabidopsis thaliana	AF489530	ATTATCACCTGAAGTTGAC	GTGTAGCAATGAGTTCTT
ADCS Arabidopsis thaliana	AY096797	TGATTGTTGACCTTCTAA	ACTGTTGTGTATGATTCT
tumor protein homologue, reference gene	LOC_Os11g43900.1	TACATCAAGAACCTCTCC	ACCAACAAAGAACTGAAG
expressed protein, reference gene	LOC_Os07g02340.1	GAACATGGAGAAGAACAA	CATATCTTGCACTGGATG
expressed protein, reference gene	LOC_Os06g47230.1	GGAGTACGACGAGAGTGAT	CCTTCTGTCCTGCCTTGG
GTPCH Irice	LOC_Os04g56710	TTATTCTGTTTCGCATTG	AACCAGTTGTATCTACTAC
DHNA isoform 1	LOC_Os09g38759	TTGCTTACAGTGACTTAA	AGGGTATATCATAGGACAT
DHNA isoform 2	LOC_Os08g44210	ATGAACAACACCTATACTTG	TCTTGACATGGATACTGAG
DHNA isoform 3	LOC_Os06g06100	CATGTCCTGTGATATAGC	GATACACTCAACAATGATATG
ADCS rice	LOC_Os06g48620.1	CCTGGATTACATAAGAGA	AGAAGAGCAACATATACT
ADCL J033051I12	LOC_Os05g15530	CTTCTAGCTGCTCTATCTG	CGATCACCAAGACATCAA
ADCL J023080D02	LOC_Os02g17330	CTCTCAGACATCTTAATCCT	TATCCACATCTTCTCTACATT
HPPK/DHPS	LOC_Os07g42632	TGAACAGAATCTACTATACG	AAGAACTATCCTCCACAA
DHFS	LOC_Os12g42870	AAGACCAGATGTTGTATT	ATATCTCCTTCAGTGATG
FPGS J033080K13	LOC_Os03g02030	AATGATACTTCTTCTTGGA	CCTGAAATTGAAATTGTTG
FPGS J023125G01	LOC_Os10g35940	GATTCAGATACTAACCAACT	GCATCAGTGACATCATAT
DHFR J090024J21	LOC_Os11g29390	CAGTGAATGATGTTAGAG	AGATGGAGACAATTAGTA
DHFR J100037K11	LOC_Os12g26060	GCCTTGTCATCCTGTCTT	ATCCATCTTATACCAATCATCCA
putative 12-oxophytodienoate reductase	LOC_Os06g11240	GCACCATTGAACAAGTATAATAG	CCCTCGTCTTTCTCATCA
WRKY19	LOC_Os05g49620	ACGGAAGGAATTCATCAG	CATAGGCTCTGCCACTAC
2Fe-2S iron-sulfur cluster binding domain containing protein	LOC_Os01g64120	CAGATCGAGCAGGGGTAC	TTCCTTGTGGGTGTAGATGA
LTPL162 - Protease inhibitor/seed storage/LTP family protein precursor	LOC_Os11g40530	GGGGAAGTCGTCACCATCAC	GCCACCATCAAGTACACAAGC
Thaumatin family domain containing protein	LOC_Os12g43450	GCGTCTTCGTCCTCCTCCTC	GTAGCGGCAGCGGTTCAC
putative VQ domain containing protein	LOC_Os10g01240	CCCTCAGACGGCGAAGTG	TGTTGAGACGGACGATTATTATGC
putative C-methyltransferase	LOC_Os03g59290	CTGATGCTGTTGGTGTCTAC	TGGTGGTTCTGATGTCTGG
Limonoid UDP-glucosyltransferase	LOC_Os02g09510	GTGTTCTCGCTCTACTACC	GGATCTGCTGCTGTATCG

Results

Folate, p-ABA and pterin analysis in transgenic rice seeds overexpressing GTPCHI and ADCS genes

In order to investigate the effect of folate enhancement on rice seed metabolism, developing seeds were sampled at 8, 12, 16, 20, 24 and 28 DPA. After complete expansion of grain length and width (around 12 DPA), endosperm development was completed around 12-16 DPA. In these early stages (corresponding to R5 and R6 reproductive developmental stages, according to Counce et al., 2000), the seeds were fully grown, green and soft. From 16 DPA on, they started to dry and harden and became yellow (20 DPA). Around 28 DPA, seeds were fully mature and brown (Supplemental Fig. S1.1). Based on the developmental folate profile in the GA 26.5 line, three time points of developing rice seeds were chosen for microarray hybridization: an early stage before folate accumulation was at the highest, the stage where folate content peaked and a later stage after this peak, when folate levels were decreased again. In order to choose the correct time points, folate content and levels of

folate precursors, p-ABA and pterins, were analyzed in developing rice seeds of WT, a homozygous G line, a homozygous A line and two GA lines; one with a high folate content (GA9.15) and one with an intermediate folate content (GA26.5) (Fig. 1.1 and Supplemental Table S1.4). Folate levels in WT rice seeds were the highest at 8 DPA (146.51 ± 17.04 μg/100 g) and decreased steadily during seed development to a level of 59.23 ± 4.27 μg/100 g at 28 DPA. A similar developmental folate profile was observed in the A line, taking into account that folate content varied greatly in the 12 DPA samples analyzed (387.74 ± 251.62 μg/100 g). Nevertheless, a small reduction was observed at 28DPA (43.01 ± 0.62 μg/100 g) as compared to the WT at the same time point. This reduction in folate content in A lines is consistent with previous reports (Storozhenko et al., 2007). In the G line, folate content increased up to 423.98 ± 2.48 μg/100 g at 16 DPA, after which it was reduced to a level of 115.87 ± 1.74 μg/100 g at 28 DPA. This represented a two-fold folate enhancement as compared to WT rice seeds at the same end stage, as previously reported (Storozhenko et al., 2007). The highest folate content was measured in the two GA lines. In GA 26.5, folate levels, already high at the earliest time point measured (1495.52 ± 617.18 μg/100 g), peaked at 12 and 16 DPA (2775.64 ± 384.85 and 2643.77 ± 78.65 μg/100 g respectively) and decreased to a level of 471.41 ± 132.50 μg/100 g. On the other hand, folate content in developing GA9.15 seeds were initially quite modest (312.09 ± 65.35 μg/100 g), but increased rapidly until a level of 2022.94 ± 60.90 μg/100g at 20 DPA, from which it decreased to 830.80 ± 239.77 μg/100 g at 28 DPA. Although folate content in the GA26.5 line at 12-16 DPA was higher than that in the GA9.15 line at 20 DPA, the difference in developmental folate profile between those two lines resulted in a later decrease of folate content and explains the higher folate content in the GA26.5 line. Wild type p-ABA levels are the highest in the early stages of rice seed development (approximately 190 μg/100 g FW; 8 DPA) and decreased slowly to a content of ± 110 μg/100 g FW at 28 DPA. On the other hand, the p-ABA content in the G-line showed the same profile as in developing wild type rice seeds, but the content itself was approximately half that of the wild type, most likely due to the enhanced production of folates. In the A line, p-ABA content stayed high throughout all time points, ranging between 500 to 600 μg/100 g FW. In both GA lines, GA26.5 and GA9.15, p-ABA content was high throughout rice seed development. In GA26.5, p-ABA content was generally higher than in GA9.15, with average levels of 353.20 μg/100 g FW and 290.79 μg/100 g FW respectively. However, the p-ABA level at 20 DPA in the GA9.15 line increased to a maximum of 515 μg/100 g FW. Pterin content in developing rice seeds of WT and A lines was generally low, with a maximum of 8.47±0.69 and 6.59±0.73 μg/100 g FW respectively at 16 DPA. As expected, the highest pterin levels were measured in the G-line. They remained high throughout all time points taken, with a maximum of 366.80±0.73 μg/100 g FW at 16 DPA. In both GA lines, pterin content was modestly enhanced at 16 DPA up to a maximum of 119.71±18.27 μg/100 g FW in GA26.5 and 95.38±6.68 μg/100 g FW in GA9.15.

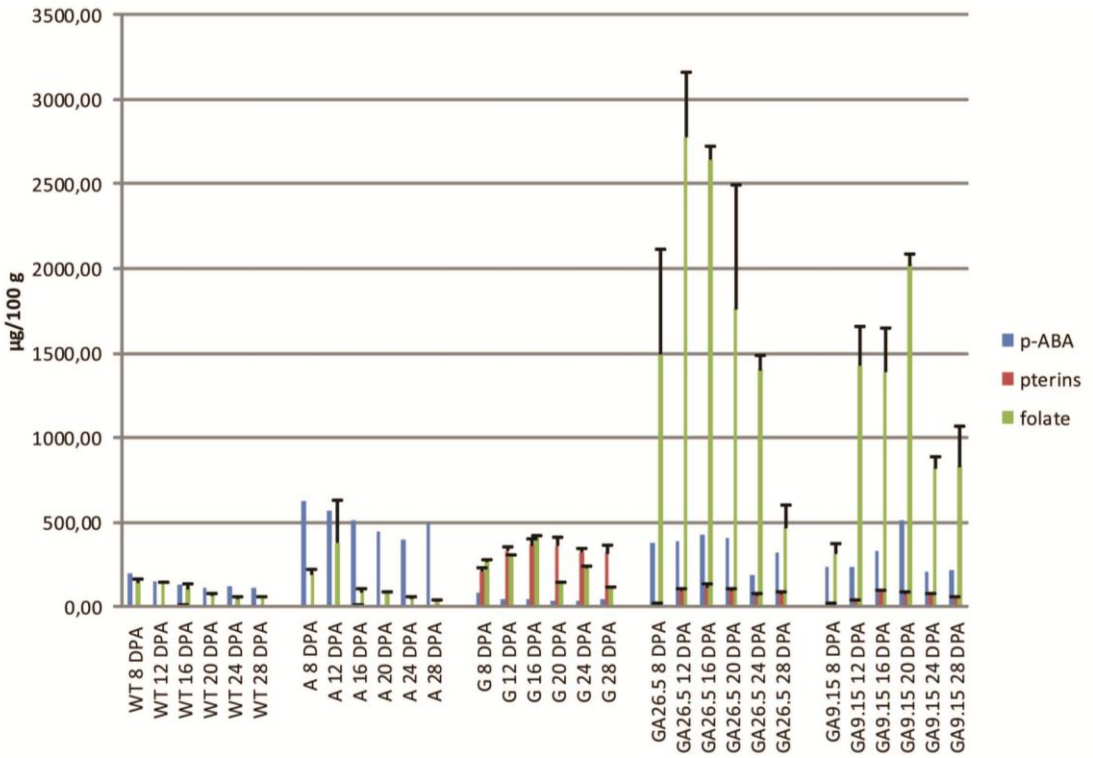


Fig. 1.1 Determination of folate, p-ABA and pterin content in developing wild type (WT) and transgenic rice lines throughout rice seed development (expressed in µg/100g). Samples were taken from 8 days post anthesis (DPA) until 28 DPA. Values are means of two independent seed samples, error bars indicate standard deviation.

Expression of *Arabidopsis* transgenes encoding GTPCHI and ADCS in developing transgenic rice seeds

To correlate folate and folate precursors profiles in developing transgenic rice seeds with the overexpression of the *Arabidopsis thaliana* (At) folate biosynthesis genes ADC synthase (*At-ADCS*) and GTP cyclohydrolase I (*At-GTPCHI*), cDNA of the five above-mentioned lines of all time points was submitted to real-time quantitative PCR analysis (Fig. 1.2). *At-ADCS* was expressed in the A line and in both GA lines. Similarly, expression of *At-GTPCHI* was detected in the G line and in the two GA lines. As expected, the transgenes were not expressed in WT rice seeds, nor in RT-minus and no template controls (data not shown). Compared to *At-GTPCHI*, *At-ADCS* expression was generally low in the developing transgenic rice seeds (with normalized expression values ranging between 0.04 ± 0.02 (GA 9.15, 28 DPA) and 2.38 ± 0.56 (GA 26.5, 12 DPA) for *At-ADCS* and between 155.35 ± 41.33 (GA 9.15, 28 DPA) and 3007.90 ± 658.75 (GA 26.5, 12 DPA) for *At-GTPCHI*). In general, *At-ADCS* transgene expression was the highest at the earliest time points taken, after which expression declined throughout rice seed development. *At-ADCS* expression dropped faster, as compared to its expression in the 2A 11.2.7.5 line, to a basal level in GA26.5 (after 12 DPA) and GA9.15 (after 8 DPA), whereas this drop occurred later in developing seeds of the 2A 11.2.7.5 line (after 16 DPA). In contrast, *At-GTPCHI* showed a peak in expression between 12 to 20 DPA in both GA lines and the G line. The expression declined rapidly after 12 DPA in the GA 26.5 line, whereas it stayed high until 20 DPA in the GA9.15 (1419.67 ± 232.06) and the 2G 17.1.5.5 line (2410.75 ± 650.83). Altogether, the

expression of both transgenes is concordant with the accumulation of folate and its precursors in the developing transgenic rice seeds.

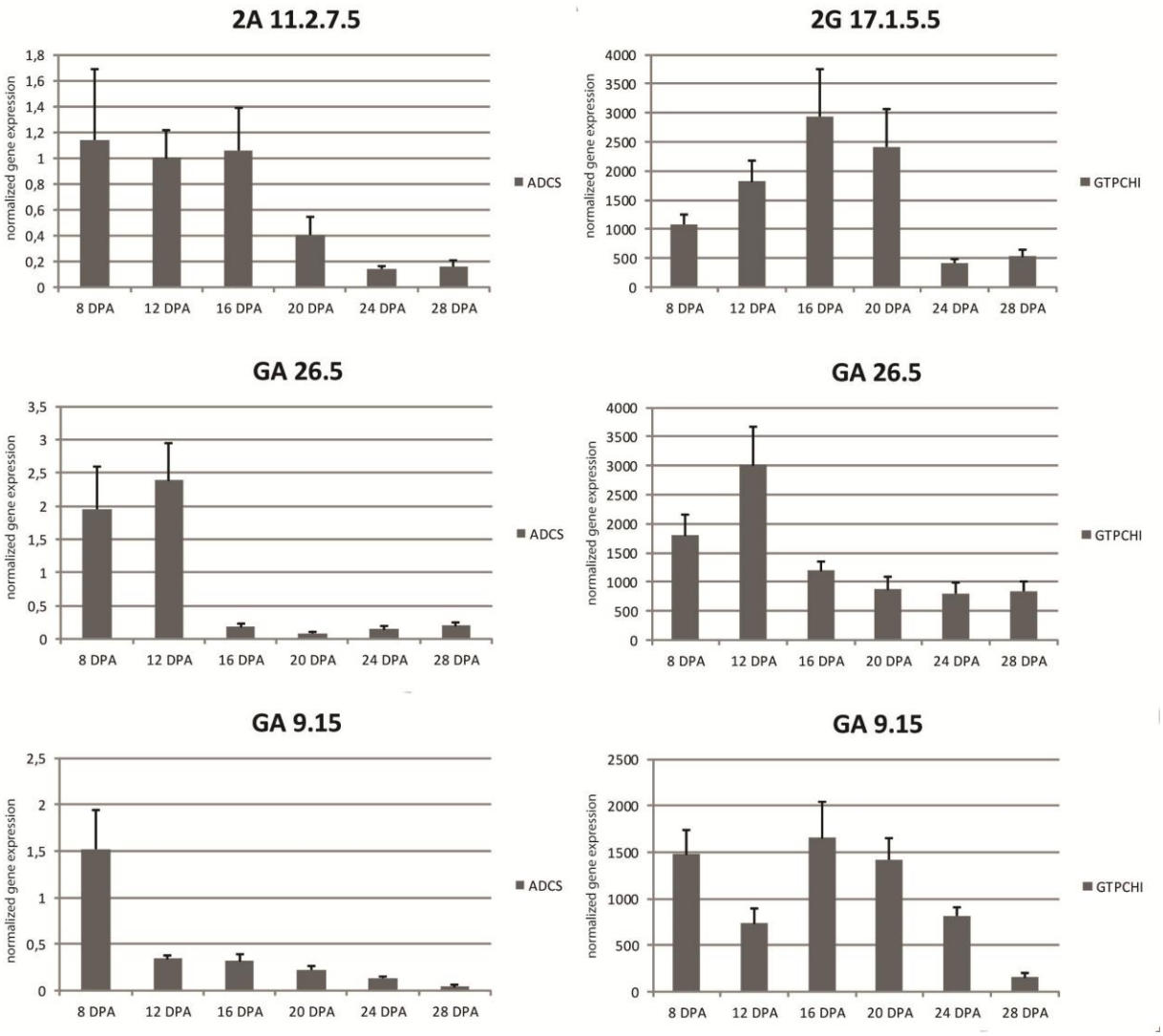


Fig. 1.2 Expression of *Arabidopsis thaliana* transgenes in four transgenic rice lines during rice seed development as assessed by real-time quantitative PCR. Three rice genes (tumor protein homologue (*LOC_Os11g43900.1)*, expressed protein (*LOC_Os07g02340.1)* and expressed protein (*LOC_Os06g47230.1*)) were used as reference genes (see Table 1) for normalization. Data analysis of three independent samples was performed using qBASE, based on the $2^{\Delta\Delta Ct}$ method. Error bars indicate standard deviation.

The effect of folate enhancement on global rice gene expression

In order to investigate the effect of folate overproduction on general rice seed metabolism, three time points from developing WT, GA26.5 and GA9.15 rice lines were chosen for a GeneChip® Rice Genome Array hybridization experiment: 8 DPA, 12 DPA and 20 DPA (Table 1.2). A total of 235 genes displayed an altered expression in one or more time points: 122 genes were up-regulated in both GA lines, 111 genes were down-regulated. When folate accumulation was the highest (12-20 DPA), a peak in the alteration of gene expression was observed: the expression of 136 genes was changed at 12 DPA (76 genes were up-regulated, 60 genes were down-regulated), while 131 genes showed an

altered expression at 20 DPA (85 genes were up-regulated, 46 genes were down-regulated). Two genes were both up- and down-regulated during rice seed development. The expression of 43 genes was affected in all three time points (34 genes up-regulated, 9 down-regulated). Using the information offered by the RiceChip Annotation Site (http://www.ricechip.org) and the Rice Genome Annotation Project (http://www. http://rice.plantbiology.msu.edu/), a function was attributed to the genes whose expression was affected in developing seeds of the GA lines. According to these functions, the genes could be classified into different categories. Although the function of most of the 235 genes with an altered expression is still unclear or putative, they could be divided into 14 functional categories. Besides 75 genes with unknown function, the three major categories which were affected are related to stress/defense response and cell death (24 genes), protein mobility/modification and degradation (23 genes) and genes involved in cell structure, cell organization and development, and the cell cycle (19 genes). Most of the genes which are involved in stress/defense response and cell death were up-regulated (20 genes were up-regulated, 4 genes were down-regulated). Interestingly, these GeneChip® Rice Genome Array hybridization data suggest that the expression of genes directly involved in folate biosynthesis and metabolism remained unaltered in folate enhanced transgenic lines. In addition, these data suggest that folate enhancement through metabolic engineering in rice endosperm has an important effect on general rice seed metabolism.

Table 1.2 Significantly regulated genes in folate-enhanced rice lines, as compared to WT, on the same DPA, based on microarray data analysis (fold change>2 of <-2, see Microarray hybridization and data analysis paragraph in Material and methods section for calculations), divided into categories by putative function. Fold changes (FC) are indicated for each time point. X= no significant change in expression. DPA; days post-anthesis.

Putative function	Locus	FC 8 DPA	FC 12 DPA	FC 20 DPA
Stress/defence response/cell death				
Leucine Rich Repeat family protein	Os11g39290	2.61	x	x
Leucine Rich Repeat family protein	Os11g42070	8.44	4.66	7.54
putative LZ-NBS-LRR class	Os11g39320	x	x	4.69
NB-ARC domain containing protein	Os12g18360	x	4.49	x
putative NBS-LRR disease resistance protein	Os06g17880	2.17	x	x
putative NBS-LRR disease resistance protein	Os11g45180	4.57	7.91	13.48
putative disease resistance protein RPM1	Os11g41170	x	x	2.16
putative disease resistance protein RPM1	Os11g12040	3.55	4.07	9.83
NB-ARC domain containing protein	Os11g39310	x	x	2.34
putative NB-ARC domain containing protein	Os11g39190	3.40	12.39	10.10
putative Non-TIR-NBS-LRR type resistance protein	Os11g42040	16.08	8.76	10.11
HSF-type DNA-binding domain containing protein	Os09g35790	x	3.11	x
putative heat shock protein DnaJ	Os03g18200	2.32	2.23	2.31
putative DnaJ homolog subfamily B member 11 precursor	Os05g06440	x	x	2.33
putative heat stress transcription factor B-1	Os04g48030	x	8.41	x
putative DnaK family protein	Os05g35400	x	x	5.65
putative Heat shock protein	Os08g39140	x	2.60	x
putative SRC2 protein	Os07g47400	3.94	2.28	x
putative DNA-binding protein DSP1	Os01g50622	x	-3.23	x
SCP-like extracellular protein	Os02g54560	x	7.00	x
Thaumatin family domain containing protein	Os12g43450	x	x	-24.44
putative stromal cell-derived factor 2-like protein precursor	Os08g34190	x	x	2.34
putative SPX domain-containing protein	Os03g29250	-3.09	x	x
Harpin-induced protein 1 domain containing protein	Os04g33990	2.43	x	x
Seed development				
putative leucine-rich repeat receptor protein kinase EXS precursor	Os02g13420	4.19	x	x
Transposons/Retrotransposons				
putative retrotransposon protein	Os11g39230	x	x	-2.59
putative transposon protein, CACTA, En/Spm sub-class	Os01g23850	x	2.83	x
putative transposon protein, CACTA, En/Spm sub-class	Os09g34290	x	-2.81	x
putative retrotransposon protein, unclassified	Os03g41200	x	-2.56	x
putative retrotransposon protein, unclassified	Os01g39136	x	2.92	x

putative retrotransposon protein, unclassified	Os10g11100	x	-2.32	x
putative retrotransposon protein, unclassified	Os12g08564	x	-3.54	x
putative retrotransposon protein, Ty1-copia subclass	Os07g18780	-3.29	-2.18	x
putative retrotransposon protein, unclassified	Os07g28850	x	x	5.80
putative transposon protein, unclassified	Os03g21660	37.58	34.83	89.63
Transport				
putative aluminum-activated malate transporter	Os06g15779	x	2.57	x
putative amino acid transporter protein	Os04g12499	x	2.02	x
putative aquaporin protein	Os04g16450	x	x	4.33
putative vesicle transport protein GOT1B	Os07g40320	x	2.68	x
putative vacuolar protein sorting-associated protein 26	Os11g41130	3.35	2.68	3.77
putative VAMP-like protein YKT62	Os01g73300	x	2.33	x
putative protein transport protein Sec61 subunit beta	Os01g38510	x	x	2.27
putative transporter family protein (galactose-proton transporter)	Os11g41870	3.09	6.84	5.28
putative Phospholipid-transporting ATPase	Os03g21680	x	x	2.62
putative solute carrier family 35 member B1	Os06g39260	x	x	2.87
putative potassium channel AKT2/3	Os05g35410	4.07	3.84	x
putative MATE efflux family protein	Os06g29844	x	x	-2.43
putative multidrug resistance-associated protein	Os01g67580	x	x	-3.30
Signal transduction				
putative 14-3-3 protein	Os11g39540	x	-2.64	-3.06
putative calmodulin binding protein	Os12g36110	x	-4.27	3.38
putative calmodulin binding protein	Os11g44630	x	x	2.46
CAMK_CAMK_like.43 - CAMK includes calcium/calmodulin depedent protein kinases	Os11g07040	x	-2.68	x
CAMK_KIN1/SNF1/Nim1_like.17 - CAMK includes calcium/calmodulin depedent protein kinases	Os03g43440	x	x	-4.88
TKL_IRAK_DUF26-lc.1 - DUF26 kinases have homology to DUF26 containing loci	Os07g35280	x	-2.39	x
putative peflin	Os12g04240	x	x	2.35
putative EF hand family protein	Os03g59770	x	2.96	2.67
putative guanylyl cyclase	Os05g37950	x	-2.73	x
putative copine-6	Os08g04130	x	x	6.57
Moving, modifying and degrading proteins				
putative RING finger protein	Os12g02210	2.65	2.78	2.99
Zinc finger, C3HC4 type domain containing protein	Os12g04590	-2.03	x	x
Zinc finger, C3HC4 type domain containing protein	Os01g16120	x	-3.94	x
putative zinc RING finger protein	Os06g03580	x	-9.42	x
putative protein binding protein	Os09g25190	-2.17	x	x
putative protein binding protein	Os09g11170	2.80	5.05	x
Protein kinase domain containing protein	Os06g43030	x	-3.58	x
putative protein kinase family protein	Os01g43350	x	-2.32	x
putative protein kinase family protein	Os10g01060	x	-2.36	x
putative protein kinase	Os03g62700	x	x	2.17
putative protein phosphotase protein	Os03g18970	x	x	3.48
OsFBDUF51 - F-box and DUF domain containing protein	Os11g36610	x	-3.58	x
OsFBDUF60 - F-box and DUF domain containing protein	Os11g38200	x	-3.02	x
OsFBX352 - F-box domain containing protein	Os10g03850	x	-2.11	x
OsFBX423 - F-box domain containing protein	Os11g33250	-2.06	x	-2.79
OsFBX424 - F-box domain containing protein	Os11g33310	-2.23	-2.27	-4.02
putative Von Willebrand factor type A domain containing protein	Os11g46000	-2.79	x	x
putative peptidase, M24 family protein	Os11g33330	x	2.69	4.11
putative peptidase, T1 family	Os07g42260	x	x	2.14
putative aspartic proteinase	Os06g40818	-3.48	x	x
putative inositol 1, 3, 4-trisphosphate 5/6-kinase	Os10g01480	x	x	2.03
putative autophagy-related protein	Os11g01010	8.70	7.84	9.61
putative ADP-ribosylation factor	Os02g47110	x	2.57	x
DNA and RNA metabolism				
putative amidophosphoribosyltransferase, chloroplast precursor	Os01g65260	-3.80	x	x
WRKY19	Os05g49620	x	-6.17	x
WRKY36	Os04g46060	-2.25	-4.04	-2.80
putative VQ domain containing protein	Os10g01240	x	-5.34	x
RNA recognition motif containing protein	Os02g03040	x	-2.34	x
OsMADS22 - MADS-box family gene with MIKCc type-box	Os02g52340	x	-2.52	x
OsMADS47 - MADS-box family gene with MIKCc type-box	Os03g08754	-2.30	x	x
putative TCP-domain protein	Os06g12230	x	-2.97	x
putative SPOC domain containing protein	Os06g39480	x	-3.16	x
RNA recognition motif containing protein	Os12g43600	-93.05	-92.52	-51.61
PHF5-like protein domain containing protein	Os11g30290	x	-2.66	x
putative pentatricopeptide	Os03g63510	x	x	4.61
B3 DNA binding domain containing protein	Os01g49830	x	-2.92	x
putative no apical meristem protein	Os11g31380	-5.92	-6.86	-5.69
putative histone-like transcription factor and archaeal histone	Os10g11580	x	x	3.00
putative MYB family transcription factor	Os01g16810	x	-2.60	x
putative leucine zipper protein-like	Os09g17810	4.27	x	x
Sugar metabolism				
putative trehalase precursor	Os10g37660	x	-2.74	-2.81
putative sucrose synthase	Os06g09450	x	x	2.13
putative alpha-amylase precursor	Os08g36910	x	6.09	x

Os3bglu8 - beta-glucosidase, exo-beta-glucansase, high similarity to Os3bglu7	Os03g49610	x	x	2.65
OsMan04 - Endo-Beta-Mannanase	Os03g61270	x	x	-3.33
Cell structure/organisation/ development/ cycle				
putative pectinesterase	Os04g54850	x	x	-2.40
putative pectinesterase	Os04g46740	x	x	-3.64
putative actin-depolymerizing factor	Os12g43340	x	x	-2.40
putative ankyrin repeat domain containing protein	Os07g46500	x	x	-2.13
putative ankyrin	Os03g47650	-2.80	x	x
putative targeting protein for Xklp2	Os11g38010	x	-2.05	x
putative receptor-like protein kinase HAIKU2 precursor	Os12g43640	-2.17	-2.27	-2.05
putative protease inhibitor/seed storage/LTP family	Os10g20860	x	-2.45	-2.21
LTPL153 - Protease inhibitor/seed storage/LTP family protein precursor	Os05g47730	x	x	4.62
LTPL162 - Protease inhibitor/seed storage/LTP family protein precursor	Os11g40530	68.60	19.61	50.45
LTPL18 - Protease inhibitor/seed storage/LTP family protein precursor	Os01g12020	-2.03	x	-2.12
putative PINHEAD	Os03g47820	x	-2.64	x
putative phytosulfokines 1 precursor	Os06g42680	x	x	3.43
PAP fibrillin family domain containing protein	Os11g38260	x	3.01	-2.60
Integral membrane protein DUF6 containing protein	Os10g14920	x	x	27.67
putative heparan-alpha-glucosaminide N-acetyltransferase	Os02g32504	x	4.62	3.43
putative endothelial differentiation-related factor 1	Os06g39240	2.28	5.07	2.75
CSLF6 - cellulose synthase-like family F; beta1,3;1,4 glucan synthase	Os08g06380	x	x	3.73
CESA7 - cellulose synthase	Os10g32980	x	x	-2.02
Hormone metabolism				
putative auxin-independent growth promoter protein	Os03g21090	x	x	-2.16
putative cytokinin dehydrogenase precursor	Os01g56810	x	-4.72	x
putative 12-oxophytodienoate reductase	Os06g11240	-18.58	-7.49	-3.05
putative gibberellin receptor GID1L2	Os11g13670	2.20	2.70	x
Redox metabolism				
putative cytochrome P450 71A1	Os02g09270	-2.74	-2.19	x
putative cytochrome P450	Os05g30890	x	x	-2.33
putative cytochrome P451	Os02g09310	-5.81	x	x
putative cytochrome P452	Os02g09330	x	x	-3.06
putative cytochrome P453	Os03g02180	x	2.83	x
putative thioredoxin family protein	Os02g56900	x	2.70	x
putative plastocyanin-like domain containing protein	Os06g36010	x	x	-3.17
putative peroxidase precursor	Os06g35480	-2.97	x	x
putative oxidoreductase, aldo/keto reductase family protein	Os04g26920	-2.95	x	x
OsPDIL2-3 protein disulfide isomerase PDIL2-3	Os09g27830	x	x	4.32
putative gamma-interferon-inducible lysosomal thiol reductase precursor	Os03g18454	x	x	2.12
FAD dependent oxidoreductase domain containing protein	Os12g43590	12.69	7.07	8.42
2Fe-2S iron-sulfur cluster binding domain containing protein	Os01g64120	x	9.51	x
Other metabolisms				
putative 1,2-dihydroxy-3-keto-5-methylthiopentene dioxygenase protein	Os10g28360	x	3.57	x
putative PME/invertase inhibitor	Os08g25070	x	x	2.45
putative O-methyltransferase	Os11g19840	x	-2.83	-7.30
putative nitrilase-associated protein	Os11g41150	2.39	2.66	x
putative aldehyde oxidase	Os10g04860	3.96	2.37	2.26
putative monoglyceride lipase	Os11g01040	x	-2.00	x
putative hydrolase protein	Os07g42140	2.42	x	x
putative hydrolase, acting on carbon-nitrogen	Os01g54530	x	x	-2.26
putative hydrolase	Os07g14600	-3.61	x	-2.28
Hemerythrin family protein	Os01g64250	x	-2.87	x
MBTB43 - Bric-a-Brac, Tramtrack, Broad Complex BTB domain with Meprin and TRAF Homology MATH domain	Os10g29050	x	-6.02	x
putative limonoid UDP-glucosyltransferase	Os02g09510	x	3.04	x
putative lectin receptor-type protein kinase	Os07g38810	2.01	2.51	3.70
putative lactate/malate dehydrogenase	Os12g43630	5.73	8.49	8.36
putative heavy-metal-associated domain-containing protein	Os11g05010	x	-2.81	x
putative glutathione S-transferase	Os01g27630	x	2.66	2.54
putative C-methyltransferase	Os03g59290	3.29	2.45	x
putative acyl-desaturase, chloroplast precursor	Os01g65830	x	x	4.00
putative chalcone synthase	Os07g17010	-2.95	x	x
putative cysteine synthase	Os06g36880	x	x	5.02
Unknown function				
Expressed protein	Os03g43100	28.15	22.12	17.72
Expressed protein	Os02g11970	x	x	7.47
Expressed protein	Os07g04330	-3.27	-2.85	-2.86
Expressed protein	Os05g01530	x	x	2.81
Expressed protein	Os02g58150	-2.31	x	-2.34
Expressed protein	Os02g37610	x	3.83	x
Expressed protein	Os07g29224	x	x	-3.69
Expressed protein	Os02g53240	5.54	7.80	5.13
Expressed protein	Os01g48820	3.51	3.81	3.99
Expressed protein	Os03g08580	x	-2.55	x
Expressed protein	Os09g09330	-4.11	x	-2.50
Expressed protein	Os05g27780	x	-4.36	x
Expressed protein	Os07g47590	x	2.16	x

Expressed protein	Os11g02290	9.31	20.17	76.38
Expressed protein	Os01g11360	-2.15	x	x
Expressed protein	Os07g29224	x	x	-2.57
Expressed protein	Os10g11310	x	3.46	x
Expressed protein	Os10g30944	x	-3.23	x
Expressed protein	Os10g04630	x	2.55	4.61
Expressed protein	Os10g39000	x	x	-2.45
Expressed protein	Os02g49640	11.93	9.64	8.71
Expressed protein	Os11g12270	20.56	13.74	16.22
Expressed protein	Os01g23880	x	x	2.74
Expressed protein	Os11g39254	x	3.07	2.23
Expressed protein	Os03g11840	x	x	-4.61
Expressed protein	Os06g44320	x	x	-4.06
Expressed protein	Os07g17560	-5.13	-2.03	x
Expressed protein	Os07g01300	9.77	3.33	2.58
Expressed protein	Os12g11690	x	x	-2.02
Expressed protein	Os09g33780	x	x	2.41
Expressed protein	Os11g12010	7.42	3.11	5.50
Expressed protein	Os07g17184	x	-2.51	x
Expressed protein	Os10g09684	-2.76	x	x
Expressed protein	Os11g39209	3.06	3.47	x
Expressed protein	Os07g28920	x	2.01	x
Expressed protein	Os05g39830	x	x	2.06
Expressed protein	Os11g39350	x	x	2.74
Expressed protein	Os03g11490	x	x	-2.01
Expressed protein	Os02g09530	-2.27	x	x
Expressed protein	Os05g44190	-2.42	-2.18	x
Expressed protein	Os06g42910	-2.69	-4.87	x
Expressed protein	Os05g50390	2.78	x	x
Expressed protein	Os05g08900	x	2.14	2.12
Expressed protein	Os08g31850	-4.34	x	x
Expressed protein	Os06g05740	x	x	2.96
Expressed protein	Os05g03934	x	2.98	x
Expressed protein	Os03g02470	136.13	28.30	18.77
Expressed protein	Os05g05390	-3.83	-2.70	-6.15
Expressed protein	Os06g36070	x	2.44	x
Expressed protein	Os08g18890	x	x	-2.20
Expressed protein	Os08g32410	-5.28	x	x
Expressed protein	Os11g42170	3.57	2.45	2.00
Expressed protein	Os12g41760	5.89	9.94	15.08
Expressed protein	Os02g01810	x	x	3.15
Expressed protein	Os05g01010	x	x	-2.34
Expressed protein	Os08g11450	x	-2.22	x
Expressed protein	Os11g42890	x	-2.40	-2.19
Expressed protein	Os11g44430	x	x	2.45
Expressed protein	Os09g11760	-2.83	x	x
Expressed protein	Os12g32580	x	2.01	x
Expressed protein	Os12g43580	-2.52	-3.58	-3.35
Expressed protein	Os07g04350	-2.25	x	x
Expressed protein	Os01g33650	-2.43	x	x
Hypothetical protein	Os09g38104	4.44	3.74	4.12
putative nodulin MtN3 family protein	Os11g31190	x	x	3.35
putative nodulin MtN3 family protein	Os03g22200	x	x	-2.03
putative nodulin	Os01g61010	x	2.09	2.52
putative DUF584 domain containing protein	Os07g33270	x	-2.28	x
putative mucin-associated surface protein	Os08g38280	2.11	2.30	2.32
putative low photochemical bleaching 1 protein	Os05g39230	x	-3.19	x
HAT dimerisation domain containing protein	Os11g39200	2.89	2.89	4.15
putative FRA10AC1	Os11g42000	2.53	2.25	2.00
putative coiled-coil domain-containing protein 25	Os01g54670	x	-6.09	x
putative circumsporozoite protein precursor	Os01g35330	31.51	24.03	31.35
putative CbbY protein-related	Os12g43520	2.73	4.22	2.67
Marker gene				
hygromycin B phosphotransferase	NO_MATCH	368.43	43.13	589.49

Validation of the microarray data by real-time quantitative PCR

To confirm the results obtained by GeneChip® Rice Genome Array microarray hybridization, real-time quantitative PCRs were conducted on all six time points of the three lines mentioned above (WT, GA 26.5 and GA 9.15). Since these folate biofortified lines also contain enhanced pterin and p-ABA levels, a G-line (2G 17.1.5.5) and an A-line (2A 11.2.7.5) were included in the analysis, to investigate

whether the altered gene expression can be attributed to folate enhancement and/or its precursors. First of all, the expression of the endogenous rice folate biosynthesis genes was investigated (Fig. 1.3). Based on data from the Rice Genome Annotation Project (http://rice.plantbiology.msu.edu/) and as previously communicated (Anukul et al., 2010), 13 putative folate biosynthesis genes were found in the rice Nipponbare genome: 1 *ADC synthase* (*ADCS*) (*LOC_Os06g48620*), 2 *ADC lyases* (*ADCL*) (*LOC_Os05g15530, LOC_Os02g17330*), 1 *GTP cyclohydrolase 1* (*GTPCHI*) (*LOC_Os04g56710*), 3 *dihydroneopterin aldolases* (*DHNA*) (*LOC_Os09g38759, LOC_Os08g44210, LOC_Os06g06100*), 1 bifunctional *dihydropterin pyrophosphokinase/dihydropteroate synthase* (*HPPK/DHPS*) (*LOC_Os07g42632*), 1 *dihydrofolate synthetase* (*DHFS*) (*LOC_Os12g42870*), 2 *dihydrofolate reductases* (*DHFR*) (*LOC_Os11g29390, LOC_Os12g26060*) and 2 *folylpolyglutamate synthetases* (*FPGS*) (*LOC_Os03g02030, LOC_Os10g35940*). Since biochemical data on enzyme activity of folate biosynthesis genes in rice is lacking, these functions remain putative. The expression of the putative rice folate biosynthetic genes was low, compared to transgene expression, at all time points and lines examined, and the expression in both transgenic GA lines followed the same trend as in the WT thus confirming the results obtained by the microarray experiment. Overall, enhancement of folate levels through metabolic engineering has no effect on the expression of the endogenous rice folate biosynthesis genes. Second, the expression of 8 randomly picked genes which showed a significantly altered expression in the microarray experiment was investigated with real-time quantitative PCR (Fig. 1.4). To evaluate whether the regulated expression was due to folate enhancement or to the accumulation of folate biosynthesis precursors p-ABA and pterines, 2G 17.1.5.5 and 2A 11.2.7.5 were included. Amongst these 8 genes, one gene is involved in stress/defense response and cell death(thaumatin family domain containing protein (*LOC_Os12g43450*)), two genes in DNA and RNA metabolism (WRKY19 (*LOC_Os05g49620*) and putative VQ-domain containing protein (*LOC_Os10g01240*)), a gene involved in cell structure, cell organization and development (LTPL 162 (*LOC_Os11g40530*)), one gene playing a role in hormone metabolism (putative 12-oxophytodienate reductase (*LOC_Os06g11240*)), one gene involved in redox metabolism (2Fe-2S iron-sulphur cluster binding domain containing protein (*LOC_Os01g64120*)), a putative C-methyltransferase (*LOC_Os03g59290*) and a putative limonoid UDP-glucosyltransferase (*LOC_Os02g09510*). In general, the data obtained by real-time quantitative PCR confirmed the microarray hybridization data. Interestingly, LTPL 162 (*LOC_Os11g40530*), clearly showed an up-regulation in expression in all time points for the two GA lines and the 2A 11.2.7.5 line, whereas it was almost not expressed in the WT and the 2G 17.1.5.5 line.

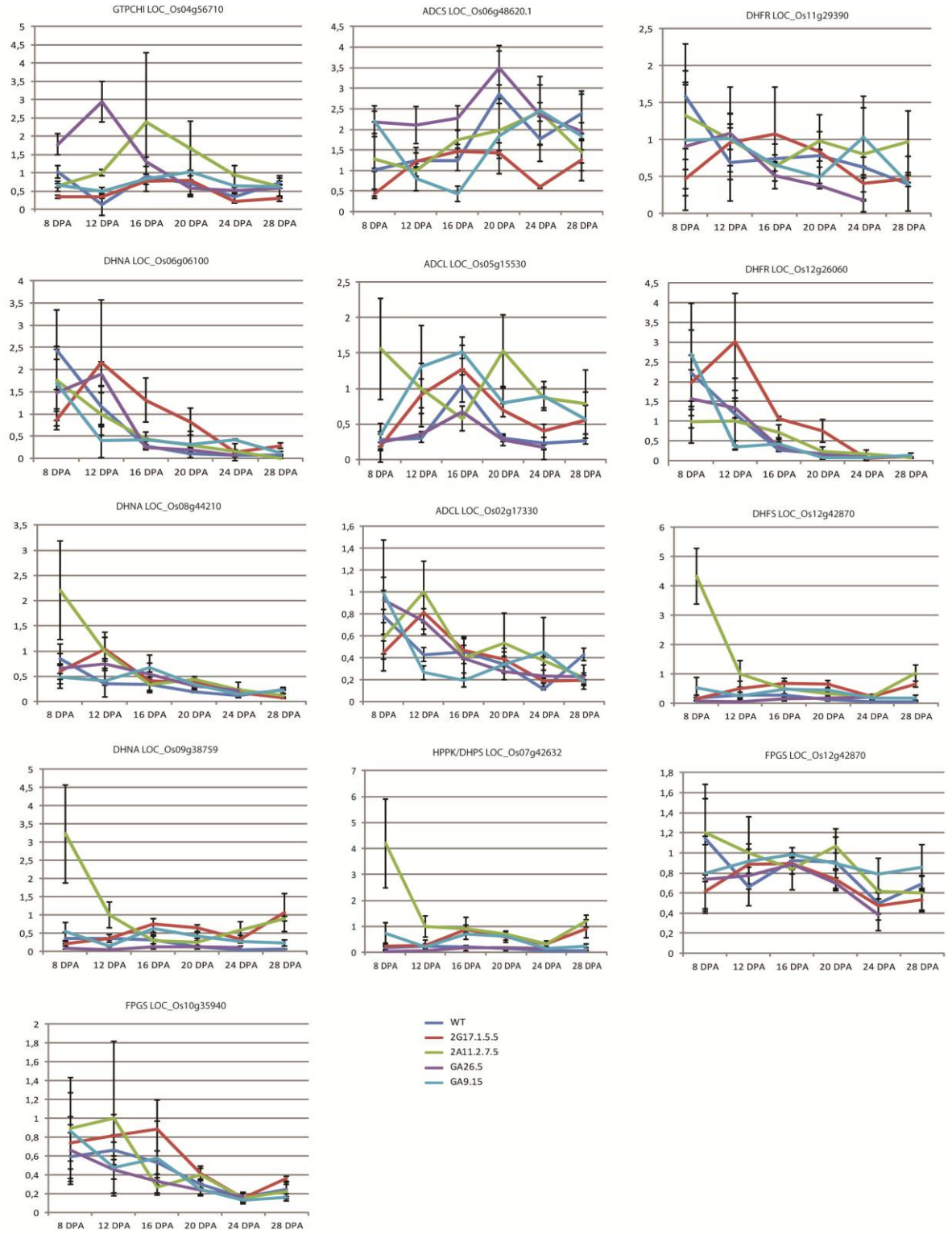


Fig. 1.3 Expression of endogenous folate biosynthesis genes during rice seed development in WT and four transgenic rice lines by real-time quantitative PCR. Three rice genes were used as reference genes (tumor protein homologue (*LOC_Os11g43900.1)*, expressed protein (*LOC_Os07g02340.1)* and expressed protein (*LOC_Os06g47230.1*)) (see Table 1) for normalization. Data analysis of three independent samples was performed using qBASE, based on the $2^{\Delta\Delta Ct}$ method. Error bars indicate standard deviation.

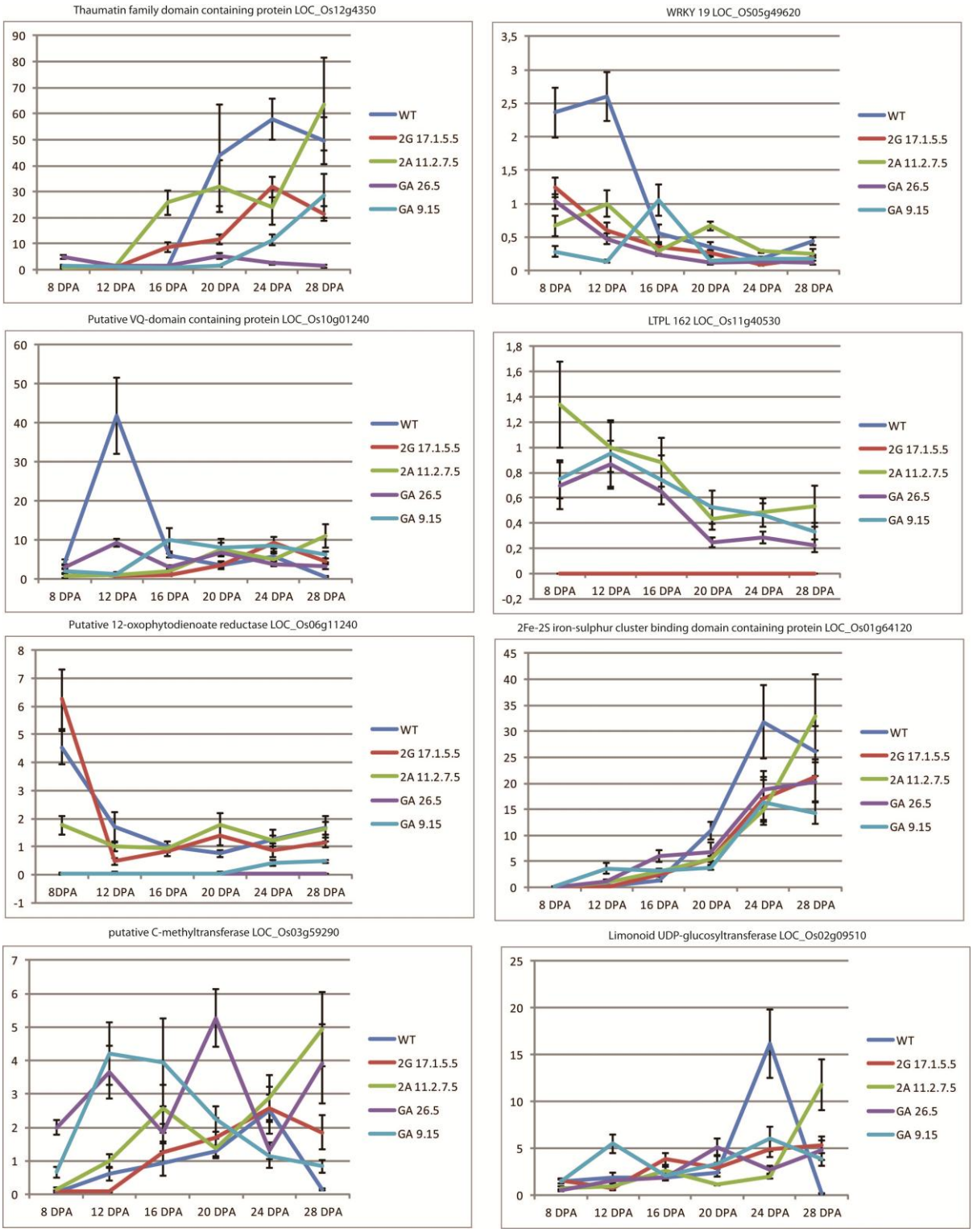


Fig. 1.4 Expression of regulated genes during rice seed development in WT and four transgenic rice lines by real-time quantitative PCR. Three rice genes were used as reference genes (tumor protein homologue (*LOC_Os11g43900.1)*, expressed protein (*LOC_Os07g02340.1)* and expressed protein (*LOC_Os06g47230.1*)) (see Table 1) for normalization. Data analysis of three independent samples was performed using qBASE, based on the $2^{\Delta\Delta Ct}$ method. Error bars indicate standard deviation.

Discussion

In this paper we describe the effects of folate enhancement on rice seed metabolism. In order to do so, folate and its precursors p-ABA and pterins were determined at 4-day intervals from day 8 to day 28 post-anthesis in 5 different lines: WT, a transgenic rice line expressing *Arabidopsis thaliana* cDNA encoding GTP cyclohydrolase I (2G 17.1.5.5), a transgenic line expressing *Arabidopsis thaliana* cDNA encoding ADC synthase (2A 11.2.7.5) and two transgenic rice lines expressing both transgenes on one T-DNA locus (GA 9.15 and GA 26.5) (Storozhenko et al., 2007) (Fig. 1.1). The transgenes are controlled by different endosperm specific promoters; GTPCHI is driven by the globulin-1 promoter, ADCS by the glutelin B1 promoter.

Folate, p-ABA and pterin analysis

In WT, folate content was the highest in the initial stages of rice seed development, when the seed is still green, from which it steadily decreases throughout the maturation and ripening of the seed. This result was as expected, since folate demand is higher in green tissue, due to its role in chlorophyll biosynthesis, photorespiration and other processes of C1 metabolism in developing tissues. A similar developmental profile was observed for p-ABA content. Since, besides folate biosynthesis, there is no other metabolic fate reported for the latter, p-ABA and folate accumulation should coincide. P-ABA accumulation in WT rice seed remained low (between 111.22 and 193.14 µg/100 g FW). On the other hand, folate content in the transgenic G-line was in the same order of magnitude of that in WT at 8 DPA (272.27±3.77 and 146.51±17.04 µg/100 g FW respectively), but raised 2.88-fold at 16 DPA (Fig. 1.1, Supplemental Table S1.4 and supplemental Fig. S1.2). Although pterin content in this transgenic line remained quite high, p-ABA levels were low (35.62 µg/100 g FW at 20 DPA), explaining the modest increase in folate content in this G line. In the transgenic A line, p-ABA levels were increased greatly at all time points taken, but folate content remained low and ended up at a lower level (43.01±0.62µg/100 g FW) than in WT rice seeds (59.23±4.27µg/100 g FW), as previously reported (Storozhenko et al., 2007). This can be explained by the low pterin content in the early stages and even a complete lack of pterins in the later stages of developing transgenic A rice seeds. In general, pterin content in the A line was lower at all time points as compared to the WT. Since the plant GTPCHI is presumed to be susceptible to feedback inhibition (Sohta et al., 1997) and the latter was proven for its mammalian counterpart (through a feedback regulatory protein (Yoneyama and Hatakeyama 1998)), a lower pterin content in developing A rice seeds could point to a regulatory mechanism driven by high p-ABA accumulation, controlling pterin biosynthesis. However, expression analysis of the endogenous pterin biosynthesis genes GTPCHI (encoding 1 isoform) and DHNA (encoding 3 isoforms) did not show a clear difference between their expression in the developing A line and in WT (Fig. 1.3). Pterins are, besides folate biosynthesis, also important in the metabolism of pigments and of several aromatic acids (Kaufman and Kaufman 1985; Kuhn and Lovenberg 1985; Johnson and Rajagopalan 1985). Thus it is possible that in developing A rice seeds, pterin content is lower due to a higher demand by these metabolic processes. Furthermore, in folate biofortified tomato, a drop in folate content in the transgenic tomato A lines could not be detected (Diaz de la Garza et al., 2007), suggesting that the mechanisms controlling pterin content in rice and tomato are different. Unfortunately, data on pterin content in transgenic A tomato was not reported, making it impossible to further substantiate this assumption. In general, folate content in both GA lines appeared to be lower at the end of maturation (471.41 ± 132.50 µg/100 g in GA 26.5 and 830.80 ± 239.77 µg/100 g in GA 9.15) as compared to the previous report (Storozhenko et al., 2007). This is

most likely caused by epigenetic effects on transgene overexpression throughout further generations of the same transgenics. Nevertheless, folate content in these GA lines is still high and they remain suitable in the battle against folate deficiency.

Expression of *Arabidopsis* transgenes encoding GTPCHI and ADCS in developing transgenic rice seeds

Folate content was high in both transgenic GA lines throughout rice seed development (Fig. 1.1). However, when comparing developmental folate profiles of GA 9.15 and GA 26.5, a clear shift was detected. The increase and peak of folate content appeared earlier in GA 26.5, the highest folate content was measured at 12 DPA (2775.64 ± 384.85 µg/100g FW), whereas this maximum was seen at 20 DPA in the GA 9.15 line (2022.94 ± 60.90 µg/100g FW). In addition, there was an earlier drop of folates in the GA 26.5 line as compared to the GA 9.15 line, resulting in a lower folate level in mature GA 26.5 seeds (28 DPA: 471.41 ± 132.50 µg/100g FW) as compared to GA 9.15 (28 DPA: 830.80 ± 239.77 µg/100g FW). To further investigate folate and folate precursor developmental profiles, the expression of the *Arabidopsis* transgenes was investigated throughout seed development of the abovementioned transgenic lines. In general, the expression of the *ADCS* transgene, controlled by the rice glutelin B1 promoter, was considerably lower as compared to the *GTPCHI* transgene (Fig. 1.2), driven by the rice globulin promoter, yet sufficient to sustain high levels of p-ABA in A and both GA lines (up to 4.5 fold higher than in the WT). The glutelin B1 promoter is considered to ensure high expression of the genes placed under its control. Indeed, glutelin is an important storage protein which accounts up to 80% of the total endosperm protein content (Yamagata et al., 1982) and transgenic rice expressing an anti-hypertensive peptide under the control of the glutelin B1 promoter, showed an accumulation of the transprotein up to 9.6% of the total seed protein (Yang et al., 2006). The accumulation of glutelin reaches a maximum at the milky stage of developing rice seeds (around 12 to 16 DPA) (Yamagata et al., 1982), which most likely coincides with a maximum expression of glutelin-associated genes. Although the lower expression of *At-ADCS*, as compared to *At-GTPCHI*, could be attributed to a natural control of the glutelin B1 promoter, it is possible that the expression of the *At-ADCS* is regulated post-transcriptionally. Little is known about such control, except that both ADC synthase and ADC lyase are not susceptible to feedback inhibition by p-ABA, the p-ABA glucose ester nor by folates (Basset et al., 2004a; Basset et al., 2004b). In summary, the accumulation of folates and its precursors in the transgenic rice lines clearly shows a close correlation with the expression of both *GTPCHI* and *ADCS* transgenes and is in line with the reported activity of their promoters (Nakase et al., 1996; Takaiwa et al., 1999).

The effect of folate enhancement on global rice gene expression

Based on the data on the developmental folate profiles in transgenic GA 26.5 rice seeds and the expression of the transgenes, three time points were chosen for a microarray hybridization experiment: 8 DPA, 12 DPA and 20 DPA. At 8 DPA, folate enhancement was still minimal in this line, whereas around 12 DPA folate accumulation peaked and folate content was much lower again at 20 DPA (Fig. 1.1). The alterations in transcriptomic activity observed (with overall changes in 235 genes in GA-lines as compared to the wild type) indicate that folate enhancement through metabolic engineering influences endogenous rice seed gene expression, as expected due to the importance of folates in C1 metabolism. Previous studies in folate biofortified tomato fruit revealed that besides 3 folate biosynthesis genes, only 14 genes showed an altered expression (Waller et al., 2010). The

TOM2 array, which was used in this transcriptomic study, represents around 11,000 unigenes (Facella et al., 2008) out of the predicted 34,727 protein-coding genes present in the tomato genome (The Tomato Genome Consortium 2012)). Due to this small gene coverage, not all genes which could possibly be affected by folate biofortification could be revealed by Waller and co-workers (Waller et al., 2010). In this study, the GeneChip® Rice Genome Array was used, which contains approximately 48,564 japonica transcripts and 1,260 indica transcripts. Taking into account that the rice genome contains more than 50,000 genes, a number which was reduced to 30,192 protein-coding genes in further optimization of its annotation (Tanaka et al., 2008), the gene coverage rate of this array can be considered close to complete. Therefore, the transcriptomic analysis performed in this study provides a more reliable view on the effects of folate engineering on plant metabolism, albeit both species-specific and tissue-specific (seeds versus fruit) differences can be expected. Indeed, even though the gene coverage in engineered tomato was smaller and therefore possibly only showed a limited effect on global gene expression in tomato fruit, comparison of the results on tomato and rice seeds clearly indicate fundamental differences between the two folate engineered species. Most strikingly, in contrast to the findings in engineered tomato, when considering both the GA 9.15 and GA 26.5 lines, no folate biosynthesis genes showed an altered expression in folate biofortified rice (Fig. 1.3 and Table 1.2). Since folates indeed accumulate in these transgenics, this indicates that the basal expression of the endogenous folate biosynthesis genes is sufficient to sustain an enhanced flux toward tetrahydrofolates. However, the rice microarray data revealed several genes with an altered expression which can be directly linked to folate metabolism. For example, iron-sulphur clusters are of great importance in gene expression, electron transfer and catalysis (Johnson et al., 2005). The biogenesis of such clusters is complex and still not fully elucidated. Recently, the importance of folates has been proven in iron-sulphur cluster formation, through the folate dependency of COG0354 proteins (Waller et al., 2012). Our microarray data show two genes with an altered expression, which are directly related to iron-sulphur clusters and thus potentially to folates: an iron-sulphur cluster binding domain containing protein (*Os01g64120*) (almost 10-fold up-regulated at 12 DPA) and a putative aldehyde oxidase (*Os10g04860*) (4-fold, 2.4-fold and 2.3-fold up-regulated at 8, 12 and 20 DPA respectively). Aldehyde oxidases belong to a multigene family (Ori et al., 1997) and catalyze the oxidation of aldehydes to form carboxylic acids. They require iron-sulphur clusters as co-factors (Coleman et al., 2002). In plants, aldehyde oxidases are necessary in the formation of abscisic acid (Leydecker et al., 1995), a plant hormone which plays an important role in seed dormancy (Hilhorst and Karssen 1992). In the transgenic GA rice, the putative aldehyde oxidase gene *Os10g04860* is up-regulated and therefore could promote seed dormancy by the production of abscisic acid.

Another gene which could be linked to folate metabolism is amidophosphoribosyltransferase (*Os01g65260*), since it encodes an important step in the *de novo* biosynthesis of purines, converting α-phosphoribosylpyrophosphate into 5-β-phosphoribosylamine (Sant et al., 1991). The action of this enzyme is allosterically inhibited by AMP, GMP, IMP and by dihydrofolate derivatives (Sant et al., 1991). The microarray data show that this gene was significantly down-regulated at 8 DPA (3.8 fold). Since folate accumulation in the transgenic GA rice seeds started early in seed development, it is possible that the accumulation of dihydrofolate derivatives could lead to the observed down-regulation , suggesting a transcriptional control, limited to the early phase in seed development (8 DPA), besides the known post-translational regulation. Since pterins are synthesized from GTP and pterin levels are higher at 12 and 20 DPA, as compared to 8 DPA (Fig. 1.1), GTP levels and thus GMP levels are expected to decrease at these time points. Consequently, the allosteric inhibition of

amidophosphoribosyltransferase by GMP would also be lower at 12 and 20 DPA, and thus concurrent with the down-regulation of gene expression at 8 DPA. Moreover, since dihydrofolate is converted to tetrahydrofolate, which accumulates throughout transgenic rice seed development, inhibition of amidophosphoribosyltransferase by dihydrofolate derivatives will decrease, gradually bringing its expression back to WT level.

C-methyltransferases belong to a large family of proteins which are able to transfer methylgroups to a series of substrates (Moore and Gaylor 1969). Since folates are of primary importance in the methylation cycle, it is likely that folate enhancement would result in an up-regulation of methyltransferases. Indeed, a putative C-methyltransferase (*Os03g59290*) was up-regulated at 8 (3.3 fold) and 12 DPA (2.5 fold) in seeds of the GA lines. Another regulated gene which could be directly linked to folate enhancement is the putative FRA10AC1 gene (*Os11g4200*). In humans, this gene is located in the vicinity of a folate sensitive fragile site and encodes a nuclear protein of unknown function (Sarafidou et al., 2004). Fragile sites are chromosomal regions where chromatin is not able to compact normally for mitosis (Sutherland 2003) and they appear as gaps or breaks in the metaphase when cells are grown under specific conditions. Although the role and significance of these fragile sites are still not completely understood, it is clear that they are regions of genomic instability, known to be correlated with specific diseases and disorders in humans, such as Jacobsen syndrome, mental retardation (borderline) and cancer (Sarafidou et al., 2004). Despite the fact that their existence has not been proven in plants, it is interesting to find that the FRA10AC1 gene showed a significant up-regulation in folate enhanced rice seeds (2.5 at 8 DPA, 2.3 at 12 DPA and 2 fold at 20 DPA). Besides the abovementioned genes directly related to folate metabolism, several genes, which are involved in rice seed development showed an altered expression upon folate enhancement. Two of these genes are involved in the control of seed size: a putative leucine-rich repeat receptor protein kinase EXS precursor (*Os02g13420*) (Canales et al., 2002) and a putative receptor-like protein kinase HAIKU2 precursor (*Os12g43640*) (Garcia et al., 2003; Garcia et al., 2005). Despite the fact that these genes were altered in expression in both GA lines, no significant difference was observed between the weight of transgenic GA, G, A and WT lines during seed development (Fig.1. 5).

Several genes related to plant stress defense appeared to be regulated upon folate biofortification of rice seeds. Plant lipid transfer proteins (LTP) are involved in somatic embryogenesis, the development of cell walls and cuticule, but also in plant defense responses against abiotic and biotic stresses (Vignols et al., 1997). In folate enhanced rice, the expression of three genes encoding LTPs, LTPL 162 (*Os11g40530*) (up-regulated 68.6, 19.7 and 50.5 fold at 8, 12 and 20 DPA respectively), LTPL 18 (*Os01g12020*) (down-regulated 2.0 and 2.1 fold at 8 and 20 DPA respectively) and LTPL 153 (*Os05g47730*) (up-regulated 4.6 fold at 20 DPA), was significantly regulated. Real-time quantitative PCR data revealed that LTPL 162 was up-regulated in both GA lines and in the A line, but not in the G line. This suggests that the altered expression could be attributed to elevated p-ABA levels, possibly pointing to a stress response, rather than to changes in seed developmental processes. In addition, a number of genes, involved in plant response upon pathogen attack, showed a significant down-regulation in the transgenic GA lines as compared to the WT control, such as a thaumatin family domain containing protein (*Os12g43450*) (down-regulated 24.4 fold at 20 DPA) and a putative VQ domain containing protein (*Os10g01240*) (down-regulated 5.3 fold at 20 DPA). Although real-time quantitative PCR confirmed down-regulation of these genes in both GA lines, as well as in the transgenic G and A line, it is possible that folate enhancement only plays an indirect role in plant defense against pathogen attack. Nevertheless, since quite a lot of genes involved in stress and

defense response and cell death were significantly up-regulated, it is clear that folate enhancement or elevated levels of its precursors evoke a stress response in the transgenic rice seeds.

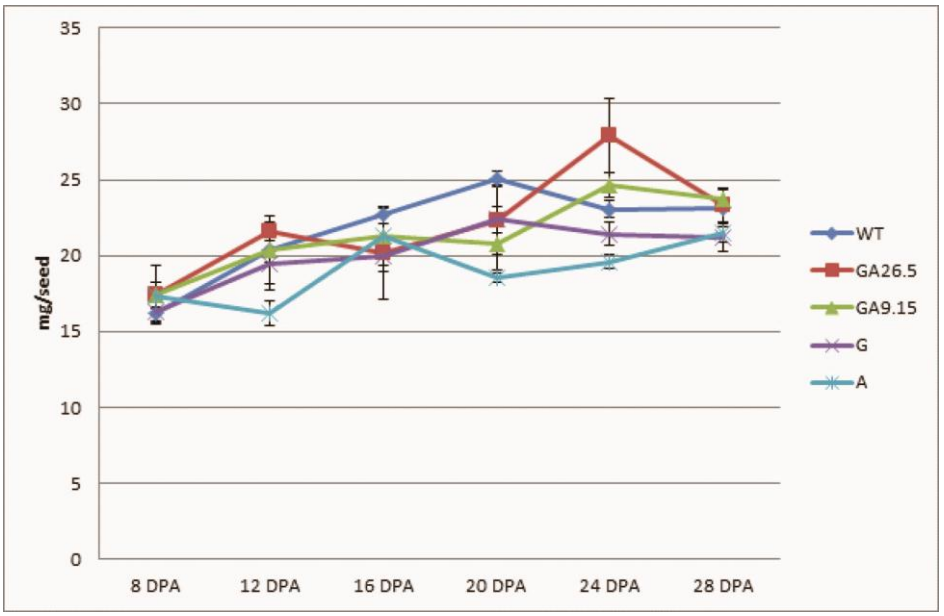


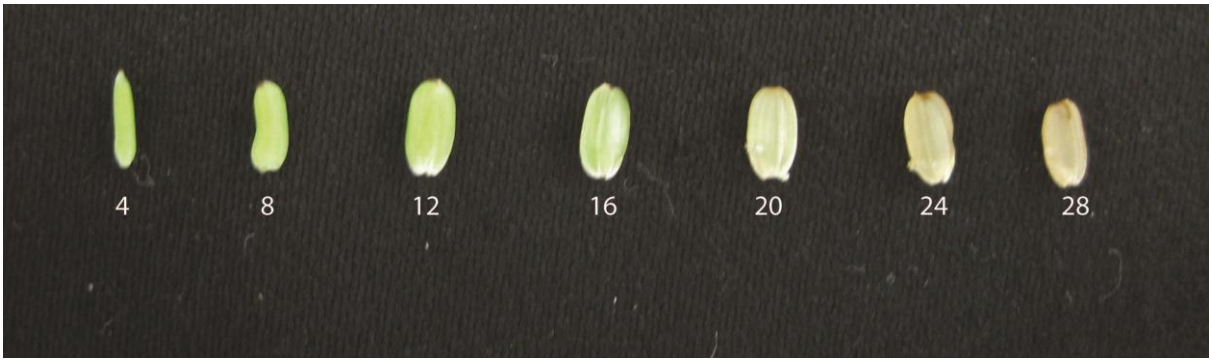
Fig. 1.5 Average weight of wild type (WT) and transgenic rice seeds during rice seed development. Calculations are based on the weight of 60 seeds from each line and time point and expressed as mg per seed. Error bars indicate standard deviations. DPA; days post-anthesis.

We conclude that folate biofortification of rice, based on seed-specific overexpression of two genes in folate biosynthesis, impacts general rice seed metabolism and affects several processes which are folate dependent, without having a significant effect on the expression of endogenous folate biosynthesis genes. The importance of folates in C1 metabolism and the high demand for folates in developing and active dividing tissues (Blancquaert et al., 2010), may explain the changes in expression of genes involved in seed development. In addition, the expression of several genes involved in plant biotic and abiotic stress and defense responses was altered. This could be attributed to folate accumulation or to the accumulation of its precursors p-ABA and/or pterins. Further optimization of the flux toward tetrahydrofolate and derivatives thereof, without accumulation of these precursors, could lower the stress burden in folate engineered rice seeds.

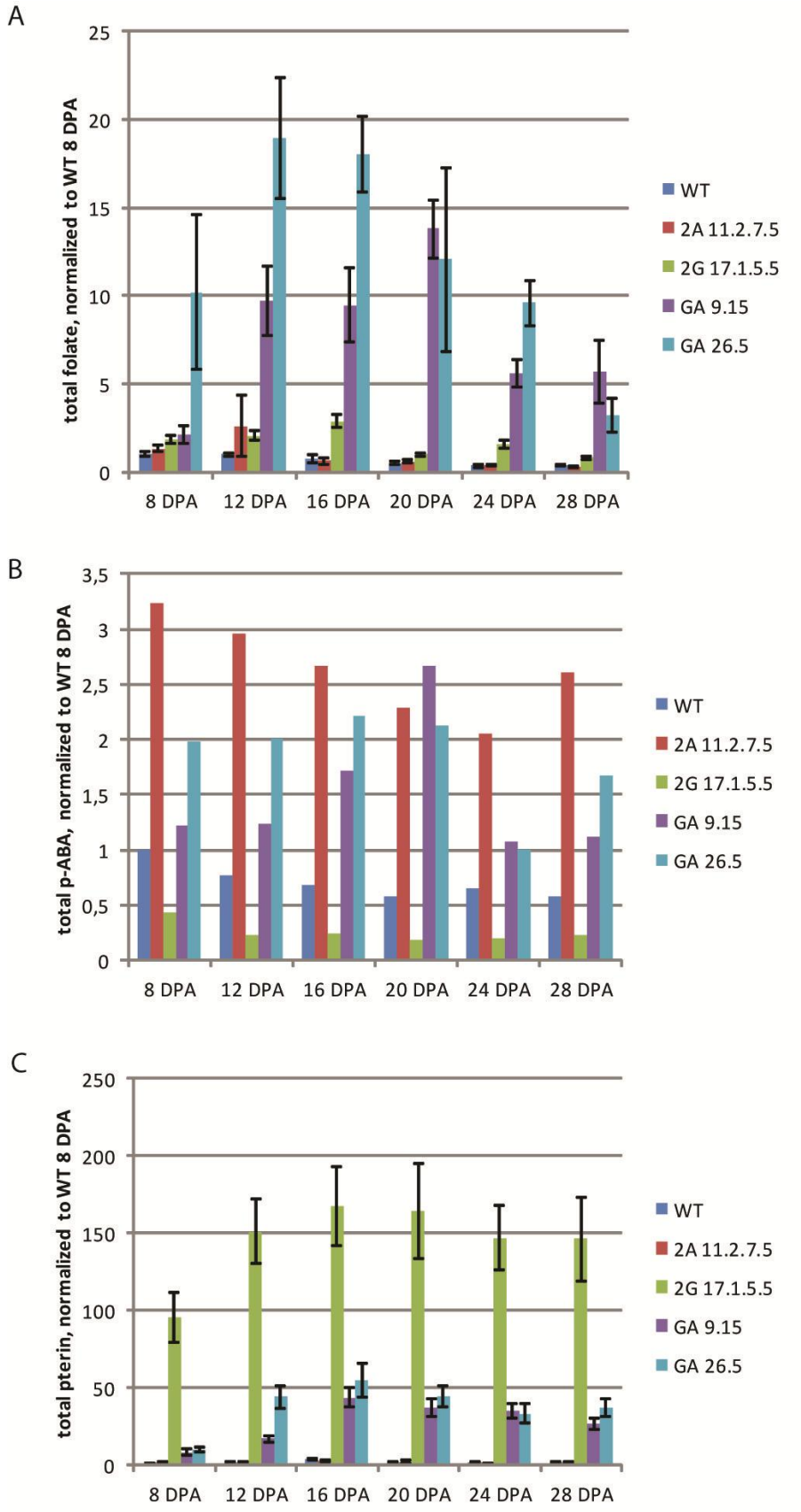
Acknowledgments

DVDS and WL gratefully acknowledge financial support from Ghent University (Bijzonder Onderzoeksfonds, BOF2004/GOA/012 and BOF2009/G0A/004), and the Fund for Scientific Research Flanders (FWO, project 3G012609).
D.B. was indebted to FWO (Fonds voor Wetenschappelijk Onderzoek – Vlaanderen), for a PhD grant.

SUPPLEMENTAL DATA



Supplemental Fig. S1.1 The different stages of rice seed development. Numbers indicate days post-anthesis (DPA). Expansion of grain length and width (around 8 DPA) is followed by the completion of the endosperm around 12-16 DPA (milky stage), when seeds are fully grown, green and soft. From 16 DPA on, the seeds started to dry and harden and became yellow (20 DPA). Around 24-28 DPA, seeds were fully maturated and dehydrated (28 DPA).



Supplemental Fig. S1.2 Total folate (A), p-ABA (B) and pterin (C) ratios in WT and transgenic 2A 11.2.7.5, 2G 17.1.5.5, GA 9.15 and GA 26.5 lines during rice seed development. Values are normalized to the content of each compound in WT at 8 DPA. Error bars indicate standard deviation.

Folate analysis	Precursor Ion (m/z)	Product Ion (m/z)	DP (V)	EP (V)	CE (V)	CXP (V)
5-MTHF	460.3	313.2	70	10	30	18
		194.2	70	10	53	14
THF	446.2	299.3	70	10	30	20
		166.4	70	10	59	12
5,10-CH^+THF	456.2	412.1	75	10	40	10
		282.1	75	10	65	16
5-CHOTHF	474.2	299.3	61	10	45	16
		166.3	61	10	52	11
FA	442.1	295.3	61	10	30	18
		176.1	61	10	57	14
10-CHOFA	470.1	295.4	65	10	35	16
		176.3	65	10	59	14
[$^{13}C_5$]5-MTHF	465.3	313.2	70	10	30	18
[$^{13}C_5$]THF	451.1	299.3	70	10	30	20
[$^{13}C_5$]5,10-CH^+THF	461.0	416.2	75	10	40	10
[$^{13}C_5$]FA	447.2	295.3	65	10	20	16

Supplemental Table S1.1 Compound parameters for folates and their internal standards. Abbreviations: DP, declustering potential; EP, entrance potential; CE, collision energy; CXP, collision cell exit potential; V, volt; 5-MTHF, 5-methyltetrahydrofolate; THF, tetrahydrofolate; 5,10-CH^+THF, 5, 10-methenyltetrahydrofolate; 5-CHOTHF, 5-formyltetrahydrofolate; FA, folic acid; 10-CHOFA, 10-formylfolic acid.

pterin analysis	Precursor Ion (m/z)	Product Ion (m/z)	DP (V)	EP (V)	CE (V)	CXP (V)
NP	252.0	146.8	-80	-10	-22	-9
		118.9	-80	-10	-34	-7
HMP	192.0	162.0	-85	-10	-22	-9
		118.9	-85	-10	-32	-5
XP	178.0	134.9	-70	-10	-24	-7
		108.0	-70	-10	-30	-5
6-FOP	190.1	146.8	-80	-10	-24	-9
		118.9	-80	-10	-30	-7
6-CAP	206.0	162.0	-70	-10	-16	-7
		118.9	-70	-10	-32	-7
15N NP	253.0	192.9	-65	-10	-20	-9
15N BP	237.4	193.1	-40	-10	-20	-5
AP	439.1	309.9	-75	-10	-30	-15

Supplemental Table S1.2 Compound parameters for pterins. Abbreviations: DP, declustering potential; EP, entrance potential; CE, collision energy; CXP, collision cell exit potential; V, volt; NP, neopterin; HMP, hydroxymethylpterin; XP, xanthopterin; 6-FOP, 6-formylpterin; 6-CAP, 6-carboxypterin; BP, biopterin; AP, aminopterin.

p-ABA analysis	Precursor Ion (m/z)	Product Ion (m/z)	DP (V)	EP (V)	CE (V)	CXP (V)
4-aminobenzoic acid (p-ABA)	138.0	77.0	50	10	30	6
3-amino-4-methylbenzoic acid (mp-ABA)	152.0	107.0	65	10	30	4

Supplemental Table S1.3 Compound parameters for p-ABA. Abbreviations: DP, declustering potential; EP, entrance potential; CE, collision energy; CXP, collision cell exit potential; V, volt.

line + time point	Folate (µg/ 100 g FW)	Pterin (µg/ 100 g FW)	p-ABA (µg/ 100 g FW)
WT 8 DPA	146.51 ± 17.04	2.19 ± 0.28	193.14
WT 12 DPA	142.29 ± 5.56	3.93 ± 1.43	148.27
WT 16 DPA	108.17 ± 29.09	8.48 ± 0.70	130.00
WT 20 DPA	72.23 ± 6.67	3.86 ± 0.50	112.67
WT 24 DPA	52.23 ± 10.72	4.07 ± 0.17	125.08
WT 28 DPA	59.23 ± 4.27	3.52 ± 0.43	111.22
A 8 DPA	195.12 ± 21.92	4.28 ± 0.68	624.27
A 12 DPA	381.74 ± 251.62	4.35 ± 0.77	571.44
A 16 DPA	89.16 ± 21.52	6.59 ± 0.73	515.62
A 20 DPA	85.52 ± 1.14	4.71 ± 1.28	440.7
A 24 DPA	54.43 ± 6.42	3.11 ± 0.00	396.02
A 28 DPA	43.01 ± 0.62	3.62 ± 0.11	503.25
G 8 DPA	272.28 ± 3.77	210.07 ± 23.46	82.86
G 12 DPA	302.28 ± 2.26	330.88 ± 18.75	45.05
G 16 DPA	422.98 ± 2.49*	366.80 ± 32±35	47.11
G 20 DPA	142.61 ± 1.87	360.85 ± 49.20	35.62
G 24 DPA	226.79 ± 15.76	322.29 ± 21.50	38.29
G 28 DPA	115.88 ± 1.75	320.52 ± 42.70	45.09
GA26.5 8 DPA	1495.52 ± 617.19	21.66 ± 2.33	381.07
GA26.5 12 DPA	2775.64 ± 384.85	96.74 ± 10.77	387.37
GA26.5 16 DPA	2643.77 ± 78.65	119.72 ± 18.27	425.96
GA26.5 20 DPA	1765.49 ± 732.54	97.61 ± 7.25	410.11
GA26.5 24 DPA	1402.43 ± 83.78	73.06 ± 9.08	192.4
GA26.5 28 DPA	471.41 ± 132.50	81.73 ± 7.49	322.24
GA9.15 8 DPA	312.09 ± 65.35	18.43 ± 3.06	235.77
GA9.15 12 DPA	1427.17 ± 234.75	36.94 ± 2.35	237.1
GA9.15 16 DPA	1388.28 ± 262.72	95.38 ± 6.68	332.34
GA9.15 20 DPA	2022.94 ± 60.90	81.69 ± 6.50	515.24
GA9.15 24 DPA	822.02 ± 64.35	77.29 ± 4.51	208.04
GA9.15 28 DPA	830.80 ± 239.77	58.42 ± 4.88	216.23

Supplemental Table S1.4 Determination of folate, p-ABA and pterin content in developing wild type (WT) and transgenic rice lines throughout rice seed development (expressed in µg/100g). Samples were taken from 8 days post-anthesis (DPA) until 28 DPA. Values are means of two independent seed samples with standard deviation. *A Student's T-test was performed as a statistical analysis to investigate whether within each line, the maximal value of each compound was significantly higher than the time point before and after (P-value<0.05).

References

Anukul N, Abilgos Ramos R, Wright AJ, Van Der Straeten D, Tucker G, Barrett DA, Bennett MJ. 2010. Relationship between total folate concentration and folate biosynthesis gene expression in rice. Proceedings in The 1st National Rice Research Conference " Moving rice research towards innovation", Kasetsart University, Thailand.

Asensi-Fabado MA, Munné-Bosch S. 2010. Vitamins in plants: occurrence, biosynthesis and antioxidant function. Trends in Plant Science **10**, 582-592.

Bailey LB. 2004. Folate and Vitamin B12 recommended intakes and status in the United States. Nutrition Reviews **62**, S14-S20.

Basset GJ, Quinlivan EP, Ravanel S, Rébeillé F, Nichols BP, Shinozaki K, Seki M, Adams-Philips LC, Giovannoni JJ, Gregory JF III, Hanson AD. 2004a. Folate synthesis in plants: the *p*-aminobenzoate branch is initiated by a bifunctional PabA-PabB protein that is targeted to plastids. Proceedings of the National Academy of Sciences of the United States of America **101**, 1496–1501.

Basset GJC, Ravanel S, Quinlivan EP, White R, Giovannoni JJ, Rébeillé F, Nichols BP, Shinozaki K, Seki M, Gregory JF III, Hanson AD. 2004b. Folate synthesis in plants: the last step of the *p*-aminobenzoate branch is catalyzed by a plastidial aminodeoxychorismate lyase. Plant Journal **40**, 453–461.

Bekaert S, Storozhenko S, Mehrshahi P, Bennett MJ, Lambert W, Gregory JF III, Schubert K, Hugenholtz J, Van Der Straeten D, Hanson AD. 2008. Folate biofortification in food plants. Trends in Plant Science **13**, 28-34.

Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society, Series B **57**, 289-300.

Besson V, Rébeillé F, Neuburger M, Douce R, Cossins EA. 1993. Effects of tetrahydrofolate polyglutamates on the kinetic parameters of serine hydroxymethyltransferase and glycine decarboxylase from pea leaf mitochondria. Biochemistry Journal **292**, 425-430.

Blancquaert D, Storozhenko S, Loizeau K, De Steur H, De Brouwer V, Viaene J, Ravanel S, Rébeillé F, Lambert W, Van Der Straeten D. 2010. Folates and folic acid: from fundamental research toward sustainable health. Critical Reviews in Plant Sciences 29, 14-35.

Blancquaert D, Navarrete O, Storozhenko S, De Steur H, Van Daele J, Dong W, Lei C, Zhang C, Stove C, Gellynck X, Viaene J, Lambert W, Van Der Straeten D. 2013. Chapter 25: Biofortified rice to fight folate deficiency. In: Preedy V (ed) Handbook of Food fortification and Health: From Concepts to Public Health Applications Springer Science, New York.

Canales C, Bhatt AM, Scott R, Dickinson H. 2002. EXS, a putative LRR receptor kinase, regulates male germline cell number and tapetal identity and promotes seed development in *Arabidopsis*. Current Biology **12**, 1718-1727.

Choi SW, Friso S. 2005. Interactions between folate and aging for carcinogenesis. Clinical Chemistry and Laboratory Medicine **43**, 1151-1157.

Coleman M, Vontas JG, Hemingway J. 2002. Molecular characterization of the amplified aldehyde oxidase from insecticide resistant *Culex quinquefasciatus*. European Journal of Biochemistry **269**, 768-779.

Cook RJ. 2001. Disruption of histidine catabolism in NEUT2 mice. Archives of Biochemistry and Biophysics **392**, 226-232.

Counce PA, Keisling TC, Mitchell AJ. 2000. A uniform, objective, and adaptive system for expressing rice development. Crop Science **40**, 436-443.

De Brouwer V, Storozhenko S, Van De Steene JC, Wille SMR, Stove CP, Van Der Straeten D, Lambert WE. 2008. Optimisation and validation of a liquid chromatography-tandem mass spectrometry method for folates in rice. Journal of Chromatography A **1215**, 125-132.

De Brouwer V, Storozhenko S, Stove CP, Van Daele J, Van Der Straeten D, Lambert WE. 2010. Ultra performance liquid chromatography tandem mass spectrometry (UPLC-MS/MS) for the sensitive determination of folates in rice. Journal of Chromatography B **878**, 509-513.

Diaz de la Garza RI, Gregory JF III, Hanson AD. 2007. Folate biofortification of tomato fruit. Proceedings of the National Academy of Sciences of the United States of America **104**, 4218-4222.

Endres M, Ahmadi M, Kruman I, Biniszkiewicz D, Meisel A, Gertz K. 2005. Folate deficiency increases postischemic brain injury. Stroke **36**, 321-325.

Facella P, Lopez L, Carbone F, Galbraith DW, Guiliano G, Perrota G. 2008. Diurnal and circadian rhythms in the tomato transcriptome and their modulation by cryptochrome photoreceptors. PLoS One **3**, e2798.

Garcia D, Saingery V, Chambrier P, Mayer U, Jurgens G, Berger F. 2003. *Arabidopsis haiku* mutants reveal new controls of seed size by endosperm. Plant Physiology **131**, 1661-1670.

Garcia D, Fitz Gerald JN, Berger F. 2005. Maternal control of integument cell elongation and zygotic control of endosperm growth are coordinated to determine seed size in *Arabidopsis*. Plant Cell **17**, 52-60.

Geisel J. 2003. Folic acid and neural tube defects in pregnancy – a review. The Journal of Perinatal and Neonatal Nursing **17**, 268-279.

Hanson AD, Roje S. 2001. One-carbon metabolism in higher plants. Annual Reviews in Plant Physiology and Plant Molecular Biology **52**, 119-137.

Hellemans J, Mortier G, De Paepe A, Speleman F, Vandesompele J. (2007) qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. Genome Biology **8**, R19.

Hilhorst HWM, Karssen CM. 1992. Seed dormancy and germination: the role of abscisic acid and gibberellins and the importance of hormone mutants. Plant Growth Regulation **11**, 225-238.

Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, Speed TP. 2003. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics **4**, 249-264.

Jiang L, Liu Y, Sun H, Han Y, Li J, Li C, Guo W, Meng H, Li S, Fan Y, Zhang C. 2013. The mitochondrial folylpolyglutamate synthetase gene is required for nitrogen utilization during early seedling development in *Arabidopsis*. Plant Physiology **161**, 971-989.

Johnson DC, Dean DR, Smith AD, Johnson MK. 2005. Structure, function, and formation of biological iron-sulfur clusters. Annual Reviews in Biochemistry **74**, 247-281.

Johnson JL, Rajagopalan KV. 1985. Pterins in other metabolic pathways. In: Blackley RL, Benkovic S (eds) Folates and Pterins, Vol 2 Wiley, New York, pp 383-399.

Kaufman S, Kaufman EE. 1985. Tyrosine hydroxylase. In: Blackley RL, Benkovic S (eds) Folates and Pterins, Vol 2 Wiley, New York, pp251-352 .

Krebs HA, Hems R, Tyler B. 1976. Regulation of folate and methionine metabolism. Biochemistry Journal **158**, 341-353.

Kuhn DM, Lovenberg W. 1985. Tryptophan hydroxylase. In: Blackley RL, Benkovic S (eds) Folates and Pterins, Vol 2 Wiley, New York, pp115-154.

Leydecker MT, Moureaux T, Kraepiel Y, Schnorr K, Caboche M. 1995. Molybdenum cofactors mutants, specifically impaired in xanthine dehydrogenase activity and abscisic acid biosynthesis, simultaneously overexpress nitrate reductase. Plant Physiology **107**, 1427-1431.

Li GM, Presnell SR, Gu LY. 2003. Folate deficiency, mismatch repair-dependent apoptosis, and human disease. The Journal of Nutritional Biochemistry **14**, 568-575.

Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta Ct}$ method. Methods **25**, 402-408.

Lohse M, Nunes-Nesi A, Krüger P, Nagel A, Hannemann J, Giorgi FM, Childs L, Osorio S, Walther D, Selbig J, Sreenivasulu N, Stitt M, Fernie AR, Usadel B. 2010. Robin: An intuitive wizard application for R-based expression microarray quality assessment and analysis. Plant Physiology **153**, 642-651.

Lucock MD, Daskalakis IG, Wild J, Anderson A, Schorah CJ, Lean MEJ, Levene MI. 1996. The influence of dietary folate and methionine on the metabolic disposition of endotoxic homocysteine. Biochemical and Molecular Medicine **59**, 104-111.

Moore JT, Gaylor JL. 1969. Isolation and purification of an S-adenosylmethionine: delta 24-sterol methyltransferase from yeast. The Journal of Biological Chemistry **244**, 6334-6340.

Nakase M, Hotta H, Adachi T, Aoki N, Nakamura R, Masumura T, Tanaka K, Matsuda T. 1996. Cloning of the rice seed alpha-globulin-encoding gene: Sequence similarity of the 5'-flanking region to those of the genes encoding wheat high-molecular-weight glutenin and barley D hordein. Gene **170**, 223-226.

Navarrete O, Van Daele J, Stove C, Lambert W, Van Der Straeten D, Storozhenko S. 2012. A folate independent role for cytosolic HPPK/DHPS upon stress in *Arabidopsis thaliana*. Phytochemistry **73**, 23-33.

Ori N, Eshed Y, Pinto P, Paran I, Zamir D, Fluhr R. 1997. TAO1, a representative of molybdenum cofactor containing hydroxylases from tomato. The Journal of Biological Chemistry **272**, 1019-1025.

Papakostas G, Cassiello CF, Iovieno N. 2012. Folates and S-adenosylmethionine for Major Depressive Disorder. Canadian Journal of Psychiatry **57**, 406-413.

Sant ME, Lyons SD, Phillips L, Christopherson RI. 1991. Antifolates induce inhibition of amido phosphoribosyltransferase in leukemia cells. The Journal of Biological Chemistry **267**, 11038-11045.

Sarafidou T, Kahl C, Martinez-Garay I, Mangelsdorf M, Gesk S, Baker E, Kokkinaki M, Talley P, Maltby EL, French L, Harder L, Hinzmann B, Nobile C, Richkind K, Finnis M, Deloukas P, Sutherland GR, Kutsche K, Moschonas NK, Siebert R, Gecz J. 2004. Folate-sensitive fragile site FRA10A is due to an expansion of a CGG repeat in a novel gene, FRA10AC1, encoding a nuclear protein. Genomics **84**, 69-81.

Scott J, Rébeillé F, Fletcher J. 2000. Folic acid and folates: the feasibility for nutritional enhancement in plant foods. Journal of the Science of Food and Agriculture **80**, 795 – 824.

Scott JM, Weir DG. 1996. Homocysteine and cardiovascular disease. The Quarterly journal of medicine Association of Physicians of Great Britain and Ireland **89**, 561-563.

Seshadri S, Beiser A, Selhub J, Jacques PF, Rosenberg IH, D'Agostino RB, Wilson PWF, Wolf PA. 2002. Plasma homocysteine as a risk factor for dementia and Alzheimer's disease. New England Journal of Medicine **346**, 476-483.

Sohta Y, Ohta T, Masada M. 1997. Purification and some properties of GTP cyclohydrolase I from spinach leafs. Bioscience, Biotechnology, and Biochemistry ***61**, 1081-1085.*

Smith GK, Mueller WT, Wasserman GF, Taylor WD, Benkovic SJ. 1980. Characterization of the enzyme complex involving the folate-requiring enzymes of *de novo* purine biosynthesis. Biochemistry **19**, 4313-4321.

Stanger O. 2004. The potential role of homocysteine in percutaneous coronary interventions (PCI): review of current evidence and plausibility of action. Cell and Molecular Biology **50**, 953-988.

Storozhenko S, De Brouwer V, Volckaert M, Navarrete O, Blancquaert D, Zhang G-F, Lambert W, Van Der Straeten D. 2007. Folate fortification of rice by metabolic engineering. Nature Biotechnology **25**, 1277-1279.

Sutherland GR. 2003. Rare fragile sites. Cytogenetic and Genome Research **100**, 77-84.

Sutherland GR, Kutsche K, Moschonas NK, Siebert R, Gecz J. 2004. Folate-sensitive fragile site FRA10A is due to an expansion of a CGG repeat in a novel gene, FRA10AC1, encoding a nuclear protein. Genomics **84**, 69-81.

Takaiwa F, Oono K, Wing D, Kato A. 1991. sequence of 3 members and expression of a new major subfamily of glutelin genes from rice. Plant Molecular Biology **17**, 875-885.

Tanaka T, Antonio BA, Kikuchi S, Matsumoto T, Nagamura Y, Numa H, Sakai H, Wu J, Itoh T, Sasaki T et al., 2008. The Rice Annotation Project Database (RAP-DB): 2008 update. Nucleic Acids Research **36**, D1028-D1033.

Tate PH, Bird AP. 1993. Effects of DNA methylation on DNA-binding proteins and gene expression. Current Opinion in Genetics and Development **3**, 226-231.

The Tomato Genome Consortium. 2012. The tomato genome sequence provides insights into fleshy fruit evolution. Nature **485**, 635-641.

Vignols F, Wigger M, Garcia-Garrido JM, Grellet F, Kader J-C, Delseny M. 1997. Rice lipid transfer protein (LTP) genes belong to a complex multigene family and are differentially regulated. Gene **195**, 177-186.

Waller JC, Akhtar TA, Lara-Nunez L, Gregory JF III, McQuinn RP, Giovannoni JJ, Hanson AD. 2010. Developmental and feedforward control of the expression of folate biosynthesis genes in tomato fruit. Molecular Plant **3**, 66-77.

Waller JC, Ellens KW, Alvarez S, Loizeau K, Ravanel S, Hanson AD. 2012. Mitochondrial and plastidial COG0354 proteins have folate-dependent functions in iron-sulphur cluster metabolism. Journal of Experimental Botany **63**, 403-411.

Yang L, Tada Y, Yamamoto MP, Zhao H, Yoshikawa M, Takaiwa F. 2006. A transgenic rice seed accumulation an anti-hypertensive peptide reduces the blood pressure of spontaneously hypertensive rats. Federation of European Biochemical Societies Letters **580**, 3315-3320.

Yoneyama T, Hatakeyama K. 1998. Decameric GTP cyclohydrolase I forms complexes with two pentameric GTPCH feedback regulatory proteins in the presence of phenylalanine or of a combination of BH4 and GTP. The Journal of Biological Chemistry **273**, 20102-20108.

Chapter 2

Enhancing pterin and para-aminobenzoate content is not sufficient to successfully biofortify potato tubers and *Arabidopsis thaliana* plants with folate

All the expression analyses were performed by Dieter Blancquaert, as well as potato in vitro and soil cultivation. Folate, p-ABA and pterin analyses were performed by Jeroen Van Daele (Faculty of Pharmaceutical Sciences, Laboratory of Toxicology). Vector cloning, *Arabidopsis* transformations and *Arabidopsis* seedling p-ABA feeding experiments were conducted by Sergei Storozhenko. Potato transformations were done by Marjan Bergervoet.

Enhancing pterin and para-aminobenzoate content is not sufficient to successfully biofortify potato tubers and *Arabidopsis thaliana* plants with folate

Dieter Blancquaert[a], Sergei Storozhenko[a], Jeroen Van Daele[b], Christophe Stove[b], Richard GF Visser[c], Willy Lambert[b], Dominique Van Der Straeten[a,#]

[a] Laboratory of Functional Plant Biology, Department of Physiology, Ghent University, K.L. Ledeganckstraat 35, 9000 Gent, Belgium

[b] Laboratory of Toxicology, Ghent University, Harelbekestraat 72, 9000 Gent, Belgium

[c] Wageningen UR Plant Breeding, Wageningen University and Research Center, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands

ABSTRACT

Folates are important cofactors in one-carbon metabolism in all living organisms. Since only plants and micro-organisms are capable of biosynthesizing folates, humans depend entirely on their diet as a folate source. Given the low folate content of several staple crop products, folate deficiency affects regions all over the world. Folate biofortification of staple crops through enhancement of pterin and *para*-aminobenzoate levels, precursors of the folate biosynthesis pathway, was reported to be successful in tomato and rice. In this study, we show that the same strategy is not sufficient to enhance folate content in potato tubers and *Arabidopsis thaliana* plants and conclude that other steps in folate biosynthesis and/or metabolism need to be engineered to result in substantial folate accumulation. Our findings provide a plausible explanation why more than half a decade after the proof of concept in rice and tomato, successful folate biofortification of other food crops, through enhancement of *p*-ABA and pterin content, has not been reported thus far.

Adapted from:

Blancquaert D, Storozhenko S, Van Daele J, Stove C, Visser RGF, Lambert W, Van Der Straeten D. 2013. Enhancing pterin and para-aminobenzoate content is not sufficient to successfully biofortify potato tubers and *Arabidopsis thaliana* plants with folate. Journal of Experimental Botany **64**, 3899-3909.

Introduction

Tetrahydrofolate and its derivatives (collectively called folates) are water-soluble B vitamins (B9), which are important players in one-carbon (C1) metabolism, where they act as cofactors which donate and accept C1 units in folate-dependent reactions (for a review see Blancquaert *et al.,* 2010). In doing so, they participate in the biosynthesis of DNA (purines and thymidylate) (Scott *et al.,* 2000), pantothenate (vitamin B5) (Asensi-Fabado and Munné-Bosch 2010), glycine and serine (Hanson and Roje 2001). Furthermore, they are involved in histidine degradation (Cook 2001) and in methionine metabolism (Hanson and Roje 2001). Folates are also of great importance in the methyl cycle and thus indirectly involved in the methylation of hormones, lipids, proteins and DNA (Scott *et al.,* 2000). Finally, in plants, folates are necessary in photorespiration, the biosynthesis of lignin, alkaloids, betaines and chlorophyll (Hanson and Roje 2001) and in N metabolism (Jiang *et al.,* 2013). They consist of three parts: a pterin moiety, a *para*-aminobenzoate (*p*-ABA) and one to fourteen γ-linked glutamate units. Folates are a family of molecules which can differ in their oxidation state, the C1 unit and the number of glutamates. Folate biosynthesis is characterized by a subcellular compartmentation of different branches of the pathway. *P*-ABA is synthesized in the plastids, whereas pterins are produced in the cytosol. Both molecules are imported into the mitochondria, where they are condensed, reduced and polyglutamylated to form tetrahydrofolate (THF) (Fig. 2.1). Only plants and micro-organisms are capable of synthesizing folates *de novo*. Therefore, humans are completely dependent on their diet as a source of folates. Green leafy vegetables, certain fruits and fermented products are rich folate sources. However, most staple crops, such as rice, maize, banana and potato contain low folate levels (USDA National Nutrient Database for Standard References; http://ndb.nal.usda.gov/). Hence, folate deficiency is a challenging problem, present in both developed and developing regions worldwide. Neural tube defects (Geisel 2003) and megaloblastic anaemia (Li *et al.,* 2003) can result from folate deficiency. Low folate intake has also been associated with a higher risk of cardiovascular and coronary diseases (Scott and Weir 1996; Stanger 2004), stroke (Endres *et al.,* 2005), major depressive disorder (Papakostas *et al.,* 2012), numerous cancers (Choi and Friso 2005) and neurodegenerative disorders (Seshadri *et al.,* 2002). Currently, several strategies are being followed to tackle folate deficiency: folic acid pill supplementation, industrial food fortification and education to change dietary habits. However, these strategies appear insufficient to fully resolve the problem of folate deficiency and often require specialized infrastructure, which is hard to implement in poorer regions of the world (where folate deficiency is often the highest). Moreover, several reports have been made on possible negative effects of folic acid (over)supplementation (Hubner *et al.,* 2007; Burdge and Lillycrop 2012). As a complementary way to fight folate deficiency, folate biofortification, the enhancement of natural folate content in food crops, has been suggested (Bekaert *et al.,* 2008). Moreover, folate biofortified crops, such as rice, have great potential to improve the health status of poorer populations suffering from folate deficiency (De Steur *et al.,* 2010). Since conventional or marker-assisted breeding is often unsuitable to obtain high folate levels, metabolic engineering of the folate biosynthesis pathway appeared to be more promising. Until now, the different attempts made to biofortify plants with folate can be divided into two categories: the overexpression of GTP cyclohydrolase 1 (GTPCHI), the first enzyme in pterin biosynthesis, and a simultaneous overexpression of both GTPCHI and aminodeoxychorismate synthase (ADCS). The single-gene strategy has been applied to *Arabidopsis thaliana* (Hossain *et al.,* 2004), tomato (Diaz de la Garza *et al.,* 2004), lettuce (Nunes *et al.,* 2009) and white corn (Naqvi *et al.,* 2009). Although pterin levels raised massively in all these attempts, only a 2 to 8.5-fold increase in

folate content could be obtained. It was suggested that a depletion of *p*-ABA would be the cause for this modest folate enhancement (Diaz de la Garza *et al.,* 2004). Therefore, the two-gene strategy was followed and massively enhanced folate levels were obtained in engineered tomato fruit (Diaz de la Garza *et al.,* 2007; up to 25-fold) and rice seeds (Storozhenko *et al.,* 2007; up to 100-fold). Potato is an important staple crop, which, although its folate content is low, represents an important source of dietary folate in many populations (Goyer and Navarre 2009). Although a four-fold variation in folate content was observed in several cultivated and wild potato species, cooked high folate potatoes could only maximally cover 28.8% of the RDA of Europeans (taking into account that average potato consumption in 2005 in Europe was 87.8 kg per capita) (Goyer and Sweek 2011). Hence, potato is a suitable candidate to biofortify folate levels through metabolic engineering. In this paper, we report on the effect of overexpression of *Arabidopsis thaliana* cDNAs encoding GTPCHI and ADCS, both under the control of tuber-specific promoters, on folate levels in Desirée potato, a main crop, red-skinned potato variety. In parallel, it was investigated whether folate enhancement could be achieved in *Arabidopsis* plants by increasing pterin and *p*-ABA levels through engineering and *p*-ABA supplementation, respectively. Our results indicate that the 2-gene strategy, previously successfully used in tomato and rice, is not sufficient for potato or *Arabidopsis*, supporting a different regulation of folate biosynthesis in these species.

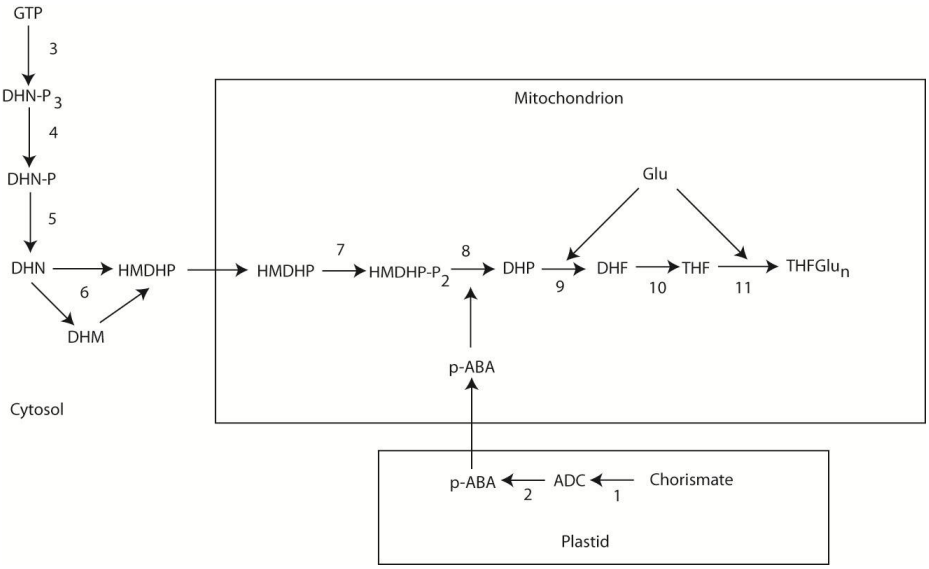


Figure 2.1: Schematic representation of the folate biosynthesis pathway in plants, which is characterized by subcellular compartmentation. Pterins are synthesized in the cytosol, whereas *p*-ABA is produced in the plastids. Both precursors are imported into the mitochondria, where synthesis of tetrahydrofolate takes place. Abbreviations: ADC, aminodeoxychorismate; GTP, guanosine triphosphate; DHN-P_3, dihydroneopterin triphosphate; DHN-P, dihydroneopterin monophosphate; DHN, dihydroneopterin; DHM, dihydromonapterin; HMDHP, 6-hydroxymethyldihydropterin; HMDHP-P_2, 6-hydroxymethyldihydropterin pyrophosphate; DHP, dihydropteroate; DHF, dihydrofolate; THF, tetrahydrofolate; $THFGlu_n$, tetrahydrofolate polyglutamate; Glu, glutamate; p-ABA, para-aminobenzoate. Folate biosynthesis enzymes: 1, ADC synthase; 2, ADC lyase; 3, GTP cyclohydrolase 1; 4, dihydroneopterin triphosphate pyrophosphatase; 5, non-specific phosphatase; 6, DHN aldolase; 7, dihydropterin pyrophosphokinase; 8, dihydropteroate synthase; 9, dihydrofolate synthetase; 10, dihydrofolate reductase; 11, folylpolyglutamate synthetase.

Materials and methods

Molecular cloning and construct design

1. Cloning *Arabidopsis thaliana* cDNAs encoding GTPCHI and ADCS into potato transformation vectors.

Three types of potato transformation vectors were constructed: vectors carrying single folate genes (*GTPCHI* and *ADCS*) under control of the patatin promoter, and a vector carrying both genes, *GTPCHI* under patatin control and *ADCS* under GBSSI control. pGKpat-fd-nos and pPGB121S plasmids, containing the patatin (PAT) and granule-bound starch synthase I (GBSSI) promoters respectively, were provided by the Wageningen UR Plant Breeding facility (Wageningen University and Research Center, The Netherlands). The patatin promoter was isolated from the pGKpat-fd-nos plasmid by restriction with *Not*I and *Nco*I. The previously created pGluB169 auxilliary vector (Storozhenko *et al.,* 2007) was cut with the same restriction enzymes to remove the glutelin B1 promoter. Upon shrimp alkaline phosphatase (SAP) treatment, the vector was ligated with the patatin promoter fragment by T4 ligase, according to the manufacturer's protocol, resulting in the pPAT69 auxilliary vector. The pPGB121S plasmid was cut with *Hind*III and *Eco*RI to isolate the GBSSI promoter, which was subsequently cloned into the auxiliary pAUX3132 vector. The resulting auxiliary vector was renamed as pGBSSI32. *Arabidopsis thaliana* transgenes *ADCS* and *GTPCHI* were amplified from the previously reported GA vector (Storozhenko *et al.,* 2007) with primers STOSER 138 and 139 and STOSER 136 and 137 respectively (all primers are listed in table S2.1). These primers contain an additional *Nco*I restriction site for GTPCHI and a *Bam*HI restriction site for ADCS. The amplicons were extracted from agarose gel, and ligated with the pJET-1 vector, to create pJET-GTPCHI and pJET-ADCS vectors. pJET-GTPCHI was cut with *Nco*I, pJET-ADCS was restricted with *Bam*HI and the resulting fragments were extracted from agarose gel. Subsequently, the above-mentioned pPAT69 and pGBSSI32 auxilliary vectors were restricted with *Nco*I and *Bam*HI respectively, treated with SAP and ligated with the GTPCHI and ADCS fragments, resulting in the pPAT-GTPCHI69 and pGBSSI-ADCS32 vectors. In order to generate the plant transformation vector, the ADCS expression cassette was removed from pGBSSI-ADCS32 by restriction with I-*Ceu*I and cloned into the pMODUL3408 vector, yielding the pGBSSI_A plant transformation vector. Likewise, the GTPCHI expression cassette was removed from pPAT-GTPCHI69 by restriction with I-*Ppo*I, and cloned into pGBSSI_A. This resulted in the creation of two plant transformation vectors: pGBSSI_A_PAT_G1 (direct orientation of PAT_GTPCHI) and pGBSSI_A_PAT_G2 (reverse orientation of PAT_GTPCHI). In order to create the pPAT_G plant transformation vector, the pGBSSI_A_PAT_G1 vector was cut with I-*Ceu*I and self-ligated to remove the GBSSI_ADCS expression cassette.

Finally, to obtain the pPAT_A plant transformation vector, the patatin promoter was amplified from the pGBSSI_A_PAT_G1 vector with the primers STOSER 178 and STOSER 179. These primers contain additional *Sac*I and *Spe*I restriction sites for traditional cloning. The amplicon was ligated with pJET-1 to create the pJET-PAT vector. Next, the 35S promoter in the pH7WG2.0 GATEWAY vector (Karimi *et al.,* 2002) was substituted by the patatin promoter. pH7WG2.0 was cut with the *Sac*I and *Spe*I restriction enzymes and ligated with the patatin promoter fragment from pJET-PAT, to form the pPAT vector. ADCS was cloned into pDONR201 by a GATEWAY BP reaction to create pADCS201 and subsequently cloned into pPAT by a GATEWAY LR reaction. This resulted in the pPAT_A plant transformation vector.

2. Cloning *Arabidopsis thaliana* cDNAs encoding GTPCHI and ADCS into *Arabidopsis* transformation vectors.

Three types of *Arabidopsis* transformation vectors were constructed: vectors carrying single folate biosynthesis genes (*GTPCHI* or *ADCS*) and a vector combining both genes. For each gene, expression was driven by the constitutive 35S promoter. In a first round of engineering*, Arabidopsis thaliana* cDNAs encoding GTPCHI and ADCS were cloned into pDONR201 by a GATEWAY BP reaction, resulting in the pGTPCHI201 and pADCS201 entry vectors (Storozhenko *et al.,* 2007). Upon GATEWAY LR reaction with pE33Gate and pE69Gate, the pGTPCHI_E33 and pADCS_E69 destination vectors were created. The ADCS expression cassette was isolated by restriction of pADCS_E69 with I-*Ppo*I, followed by gel extraction. In parallel, pMODUL3408, carrying the 35S promoter, was cut with the same restriction enzyme, treated with SAP and purified from an agarose gel. The ADCS expression cassette and the linearized pMODUL3408 vector were ligated, which resulted in the pATH_A plant transformation vector. The GTPCHI expression cassette was isolated by restriction of pGTPCHI_E33 with PI-*Psp*I, followed by gel extraction. The cassette was subsequently cloned into the pATH_A plant transformation vector. This yielded the pATH_GA plant transformation vector. Finally, the ADCS expression cassette was removed by restriction with I-*Ppo*I and self-ligation, giving rise to the pATH_G plant transformation vector.

In a second round of engineering, *Arabidopsis thaliana* cDNA encoding ADCS was cloned into pDONR221 by a GATEWAY BP reaction, which resulted in the pADCS221 entry vector. The pADCSH7 plant transformation vector was created upon LR reaction between pADCS221 and pH7WG2.0 (Karimi *et al.,* 2002). pADCSH7 carried the *hygromycin phosphotransferase II* as a selectable marker gene and the 35S promoter to drive ADCS expression.

Microbial strains, plant material and transformations

Escherichia coli strain DH-5α was used for cloning, plasmid manipulations and propagation of the Gateway™ vectors used for potato and *Arabidopsis thaliana* transformations. T-DNA from the binary vectors was delivered to potato and *Arabidopsis thaliana* cv. Columbia-0 (Col-0) cells by *Agrobacterium tumefaciens* strain LBA4404. Potato (*Solanum tuberosum* L.) variety Desirée was propagated *in vitro* on Murashige and Skoog (MS) medium by cuttings and grown in long day light conditions (16h of light, 8 h of darkness, 20°C) prior and after transformations. Transformed potato plants were transferred to soil and grown in long day light conditions (20°C). The same conditions were used for the growth of transgenic *Arabidopsis thaliana* plants. Potato transformations were performed according to (Filati *et al.,* 1987) and (Hoekema *et al.,* 1989), as adapted by (Pel *et al.,* 2009). Potato plants were selected on kanamycin (Kan) and hygromycin (Hyg) containing solid media, upon transformation with pPAT_G, pGBSSI_A, pGBSSI_A_PAT_G1, and pGBSSI_A_PAT_G2 (Kan) and pPAT_A (Hyg), respectively. Genomic DNA was isolated from potato leaves using the Invisorb® Spin Plant Mini Kit (Invitek), according to the manufacturer's protocol. Presence of the transgenes was confirmed by PCR amplification using STOSER 28 as forward and STOSER 29 as reverse primer for ADCS and STOSER 42 as forward and STOSER 43 as reverse primer for GTPCHI. *Arabidopsis thaliana* transformation was mediated by the 'floral-dip' method (Clough and Bent 1998). The cloned vector was introduced into *Agrobacterium tumefaciens* by electroporation (Gene Pulser Xcell, Bio-Rad) according to the manufacturer's protocol. Selection of transformed *Arabidopsis* plants was performed on MS solid media containing kanamycin (for plants transformed with pATH_A1, pATH_G and pATH_GA vectors) and hygromycin (for plants transformed with the pADCSH7 vector) and

transferred to soil for further growth and seed production. Transgenic *Arabidopsis* plants were selected on single copy transformation events in T_2 generation by 3:1 segregation of the kanamycin and hygromycin resistance markers and subsequently selected on homozygosity of the transgenes in T_3. Genomic DNA was isolated using the Invisorb® Spin Plant Mini Kit (Invitek), according to the manufacturer's protocol. Presence of the transgene in the genome of *Arabidopsis* plants, transformed with pADCSH7 was confirmed by PCR amplification with STOSER 39 and STOSER 73 primers for ADCS. These primers are intron-overspanning to distinguish the endogenous *Arabidopsis* gene from the transgene.

RNA extraction, cDNA synthesis and real-time quantitative PCR

Two to three freshly harvested potato tubers from each transgenic line were peeled and cut into smaller pieces. Total RNA from 100 mg of this potato tuber tissue mixture was extracted using Trizol™ reagent (Invitrogen), following the manufacturer's protocol, with minor modifications, and further purified with the aid of the RNeasy Plant Mini Kit (Qiagen). Samples were treated with DNase (RNase-free DNase set, Qiagen). Subsequently, 1 µg of potato RNA was converted into cDNA with the Verso cDNA Kit (Thermo-Scientific) with oligo-dT primers. Similarly, total RNA was extracted from 100mg of fresh *Arabidopsis* leaf tissue with the RNeasy Plant Mini Kit (Qiagen), following the manufacturer's instructions with minor modifications. The resulting RNA samples were treated with DNase (DNase I Amplification Grade Kit, Invitrogen) to remove residual DNA. cDNA was produced by reverse-transcription of the RNA samples as described for the potato samples. To investigate the expression of the transgenes and the endogenous folate biosynthesis genes encoding ADCS and GTPCHI in potato, real-time quantitative PCR was performed on the potato cDNA samples. As reference genes, potato housekeeping genes *elongation factor 1-α* (*ef1α*) and *cytoplasmic ribosomal protein L2* (*L2*) were used (Nicot *et al.,* 2005). STOSER 150 and STOSER 151 were used as forward and reverse primer for ef1α and STOSER 152 and STOSER 153 as forward and reverse primer for L2. The expression of the transgenes was investigated by using STOSER 154 as a forward primer and STOSER 155 as a reverse primer for GTPCHI and STOSER 156 as a forward primer and STOSER 157 as a reverse primer for ADCS. Endogenous potato genes expression was measured by using pot GTPCHI 1F as a forward primer and pot GTPCHI 1R as a reverse primer for GTPCHI and pot ADCS 1F as a forward primer and pot ADCS 1R as a reverse primer for ADCS. cDNA samples were diluted 5 times with deionized water and real-time quantitative PCR was conducted using the iCycler (Bio-Rad) as a thermal cycler and IQ™5 (Bio-Rad) as the optical module for Sybr Green signal detection, in a 25 µl reaction mixture composed of 5 µl diluted cDNA, 800nM of each primer and the MAXIMA™ SYBR Green qPCR Master Mix (Fermentas). All reactions included a technical replicate. A real-time quantitative PCR run comprised 40 cycles (3 min of initial denaturation at 95°C, 15 sec of denaturation at 95°C, 30 sec of annealing at 55°C, 30 sec of extension at 72°C, ended by a reaction conservation at 4°C). Data analysis and normalization was performed using the qBASE software, based on the $2^{-\Delta\Delta Ct}$ method (Hellemans *et al.,* 2007; Livak *et al.,* 2001). To investigate the expression of GTPCHI and ADCS in transgenic *Arabidopsis thaliana* lines and WT, the same conditions and materials were used as for the expression analysis of the potato samples. Here, the housekeeping genes *actin11* and *cytosolic glyceraldehyde-3-phosphate dehydrogenase* (*GapC*) were selected as reference genes for normalization with ACTIN11 F and GapC F as forward primers and ACTIN11 R and GapC R as reverse primers. Data analysis and normalization were as described for the potato samples.

p-ABA feeding experiments on Arabidopsis seedlings

For the chronic exposure of *Arabidopsis* seedlings with *p*-ABA, seeds were germinated and grown for two weeks on half-strength MS media supplemented with 100 µM of *p*-ABA or deionized water. The seedlings were collected and flash-frozen in liquid nitrogen prior to *p*-ABA and folate analysis. For the one time exposure, two week old *Arabidopsis* seedlings were vacuum-infiltrated with 100 µM *p*-ABA or deionized water. The seedlings were collected after four hours of exposure, flash-frozen in liquid nitrogen and *p*-ABA and folate content was determined.

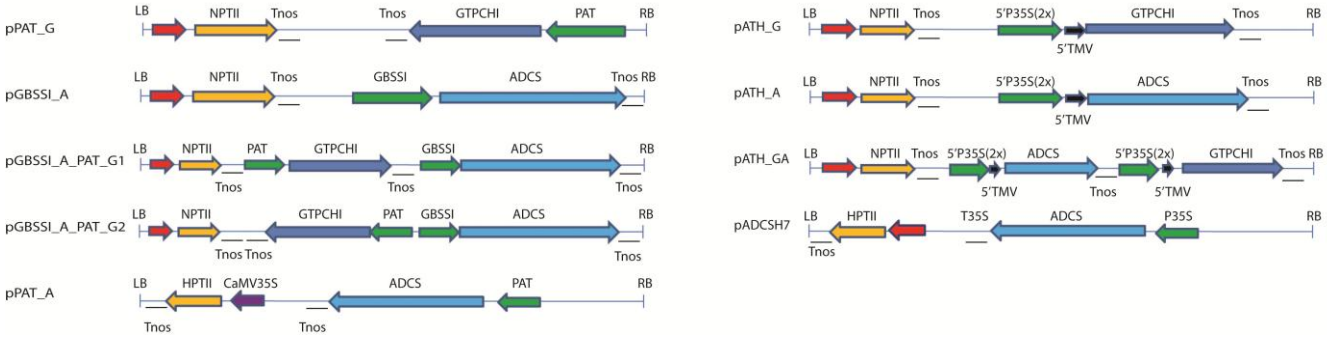


Figure 2.2: Schematic representation of the T-DNA vectors used for potato (left) and *Arabidopsis* (right) transformation. Abbreviations: LB and RB, left and right T-DNA borders respectively; NOS promoter, nopaline synthase promoter (red arrows); Tnos, nopaline synthase transcriptional terminator; PAT, potato patatin promoter; GBSSI, potato granule-bound starch synthase I promoter; 5'TMV, 5' untranslated region of tobacco mosaic virus gene; CaMV35S2 and 5'P35S(2x), core cauliflower mosaic virus 35S promoter with duplicated enhancer sequence; NPTII, neomycin phosphotransferase II; HPTII, hygromycin B phosphotransferase II; GTPCHI, *Arabidopsis thaliana* cDNA encoding GTP cyclohydrolase I (dark blue arrows); ADCS, *Arabidopsis thaliana* cDNA encoding aminodeoxychorismate synthase (light blue arrows).

Results

Transformation of potato with GTPCHI under the control of the patatin promoter and ADCS under the control of the granule-bound starch synthase I promoter

In a first round of metabolic engineering, potato (*Solanum tuberosum* L.) var. Desirée was stably transformed with *Arabidopsis thaliana* cDNA encoding GTP cyclohydrolase I (GTPCHI) under the control of the potato tuber-specific patatin (PAT) promoter (Naumkina *et al.,* 2007; Romanov *et al.,* 2007) and *Arabidopsis thaliana* cDNA encoding aminodeoxychorismate synthase (ADCS), controlled by the potato tuber-specific granule-bound starch synthase I (GBSSI) promoter (van der Steege *et al.,* 1992). The plant transformation vectors used for potato transformations are shown in Fig. 2.2. Four vectors were used: pPAT_G, pGBSSI_A, pGBSSI_A_PAT_G1 (direct orientation of GTPCHI, driven by the patatin promoter) and pGBSSI_A_PAT_G2 (reverse orientation of the GTPCHI expression cassette). These transformations resulted in G, A, GA1 and GA2 potato lines respectively. Transgenic and wild type (WT) potato plants were grown in soil and young tubers of WT, 13 G lines, 5 A lines, 6 GA1 lines and 13 GA2 lines were analyzed on folate content (Fig. 2.3A). Folate concentration ranged between 21.96 ± 1.07 µg/100 g of fresh weight (FW) (in A6) and 51.78 ± 2.02 µg/100 g FW (in GA2-7),

which represents a low enhancement in some transgenic potato lines as compared to the WT control. Expression analysis revealed that although *Arabidopsis* GTPCHI expression was high in most lines examined, the expression of *Arabidopsis* ADCS was almost undetectable (Fig. S2.1). Pterin content of 9 potato G lines and WT was measured (Fig. 3B). Pterin levels ranged between 22.68 ± 10.29 µg/100 g FW (in WT) and 288.45 ± 56.49 µg/100 g FW (in G16) and in most cases, a good correlation could be found between high pterin accumulation and *Arabidopsis* GTPCHI expression (Fig. 2.3C).

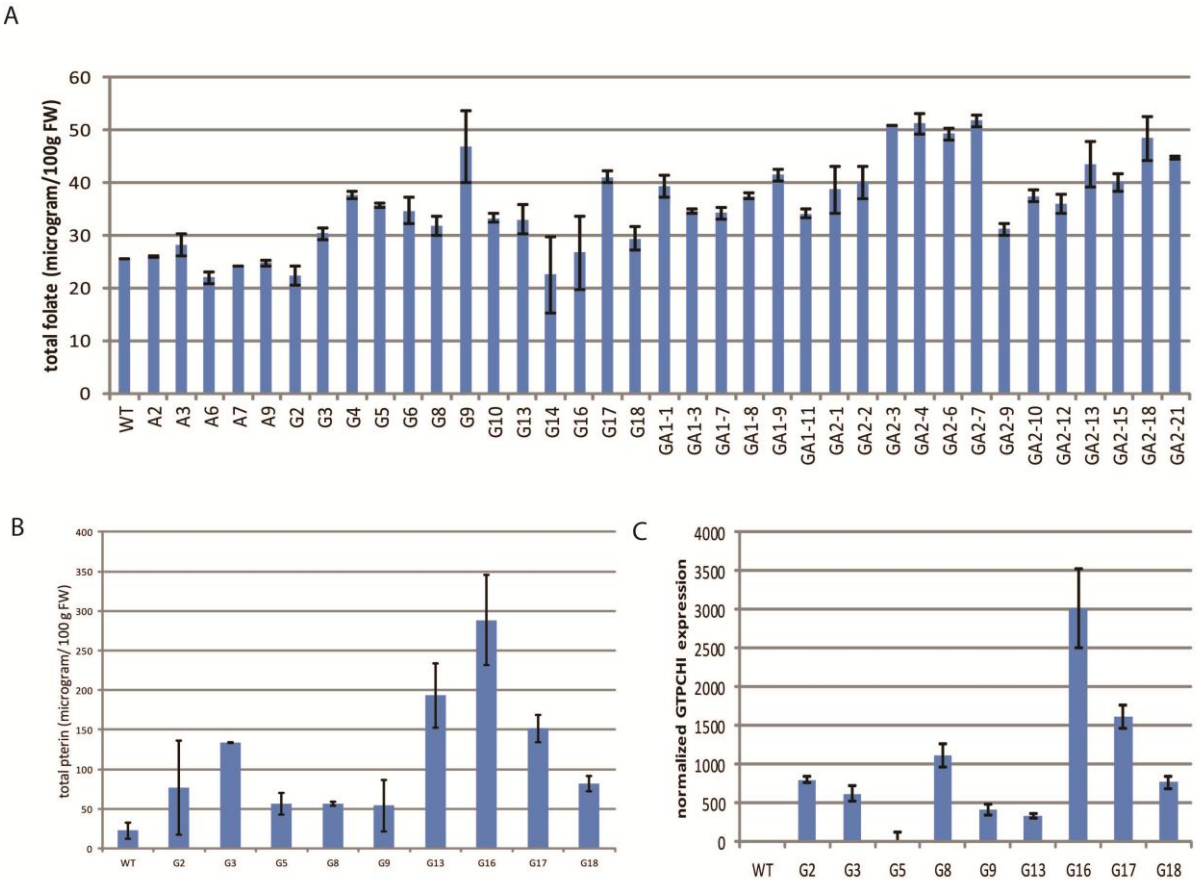


Figure 2.3: Folate, pterin and transgene expression levels from potato lines transformed with pPAT_G (G lines), pGBSSI_A (A lines), pGBSSI_A_PAT_G1 (GA1 lines) and pGBSSI_A_PAT_G2 (GA2 lines). A: Total folate content of WT, G, A, GA1 and GA2 potato tubers, expressed in µg/100 g of fresh weight (FW). Values are averages of 2 independent samples, error bars indicate standard deviation. B: Total pterin content in WT and 9 G lines, expressed in µg/100 g of (FW). Values are averages of 2 independent measurements, error bars indicate standard deviation. C: Expression of *Arabidopsis thaliana* GTPCHI in WT and 9 G potato lines by real-time quantitative PCR. Potato cytoplasmic ribosomal protein L2 was used as a reference gene for normalization. Data analysis of two independent samples was performed using qBASE, based on the $2^{-\Delta\Delta Ct}$ method. Error bars indicate standard deviation.

Retransformation of WT and pterin biofortified potato with ADC synthase under the control of the tuber-specific patatin promoter

Since expression of the ADCS transgene could not be detected in the first generation of transgenic potato, two G lines with a high pterin content and a good GTPCHI transgene expression were selected for retransformation with a new ADCS transformation vector. Since the patatin promoter

yielded satisfying results for GTPCHI transgene overexpression, the same promoter was used to control tuber-specific ADCS expression. Potato G16 and G17 lines, together with a WT control, were retransformed with pPAT_A (Fig. 2.2). The new transformants were grown in soil, together with their parental lines and WT. Tissue from young micro tubers (2 to 3 cm in diameter) was sampled and analyzed for folate, pterin and *p*-ABA content (Fig. 2.4). Retransformants were renamed as GA16-x lines and GA17-x lines, with G16 and G17 parental background respectively. Folate content ranged between 42.45 ± 1.13 (A6) and 122.68 ± 5.03 µg/100 g FW (GA17-29) (Fig. 2.4A; wild type level= 54.37 ± 0.69 µg/100 g FW). Since high pterin (Fig. 2.4B) and *p*-ABA (Fig. 2.4C) levels could be measured in most samples from both GA16 and GA17 lines (with the highest *p*-ABA content found in GA17 lines), this modest, up to 3-fold, increase in folate content, as compared to the WT control, did not result from *p*-ABA nor pterin depletion. Next, the expression of transgenes was investigated in the new transformants (Fig. 2.5A) by real-time quantitative PCR. A good expression of the ADCS transgene was found in most lines examined, especially in the GA17 lines, although it was generally lower than GTPCHI transgene expression. To investigate whether *Arabidopsis* ADCS and GTPCHI overexpression had an impact on the expression of the corresponding endogenous potato genes (potato GTPCHI: sequence IDPGSC0003DMC400034874 and potato ADCS: sequence ID PGSC0003DMC400017266) , a real-time quantitative PCR was conducted on cDNA of the same micro tuber potato samples (Fig. 2.5B). The expression of both endogenous potato genes was generally low and no clear correlation could be found between the expression of the endogenous genes and the transgenes, nor with the content of folate and its precursors. Interestingly, the expression of endogenous ADCS was generally higher than that of GTPCHI. To investigate whether folate enhancement would occur at a later developmental stage of the transgenic potato tubers, folate levels were analyzed in mature potatoes (Fig. S2.3A) from four GA16 and 5 GA17 lines. Although a good expression of both transgenes could still be detected (Fig. S2.2), folate content was lower in mature potatoes as compared to folate levels in micro tubers and no difference in folate content was detected between the transgenics and WT. The predominant folate form in all lines examined was 5-methyltetrahydrofolate, which represented approximately 50% of the total folate pool (Fig. S2.3B).

Constitutive overexpression of GTPCHI and ADCS in Arabidopsis thaliana

Two attempts were made to enhance folate content in *Arabidopsis thaliana*. In a first round of engineering, WT *Arabidopsis* was transformed with the pATH_A (ADCS driven by the 35S promoter; A lines), the pATH_G (GTPCHI driven by the 35S promoter; G lines) and the pATH_GA (GTPCHI and ADCS, both driven by the 35S promoter; GA lines) plant transformation vectors (Fig. 2.2). Expression analysis on 23 primary transformants (6 A lines, 8 G lines and 9 GA lines) showed overexpression (more than five-fold versus the GTPCHI expression in WT) of GTPCHI in 8 transgenic lines (6 G lines and 2 GA lines) (Fig. S2.4). However, no ADCS overexpression could be detected and the folate content did not differ from that in WT *Arabidopsis* (data not shown). Pterin levels in the G lines increased up to 439.03 ± 130.92 µg/100 g FW as compared to 24.49 ± 6.62 µg/100 g FW measured in the empty vector control. In a second round of engineering, 6 homozygous G lines (G1-5, G2-1, G3-4, G5-1, G6-4 and G9-3) and an empty vector control line (Ct1-6-1) were retransformed with the pADCSH7 (again with ADCS driven by the 35S promoter) plant transformation vector (Fig. 2.2). A total of 67 primary transformants were obtained and leaf samples of 60 transformants were analyzed for ADCS expression by real-time quantitative PCR. Despite the fact that five G+A lines showed an ADCS expression level 3-5 times higher than that of the empty vector control line, the folate content in none of those transgenics was significantly enhanced (Fig. S2.6).

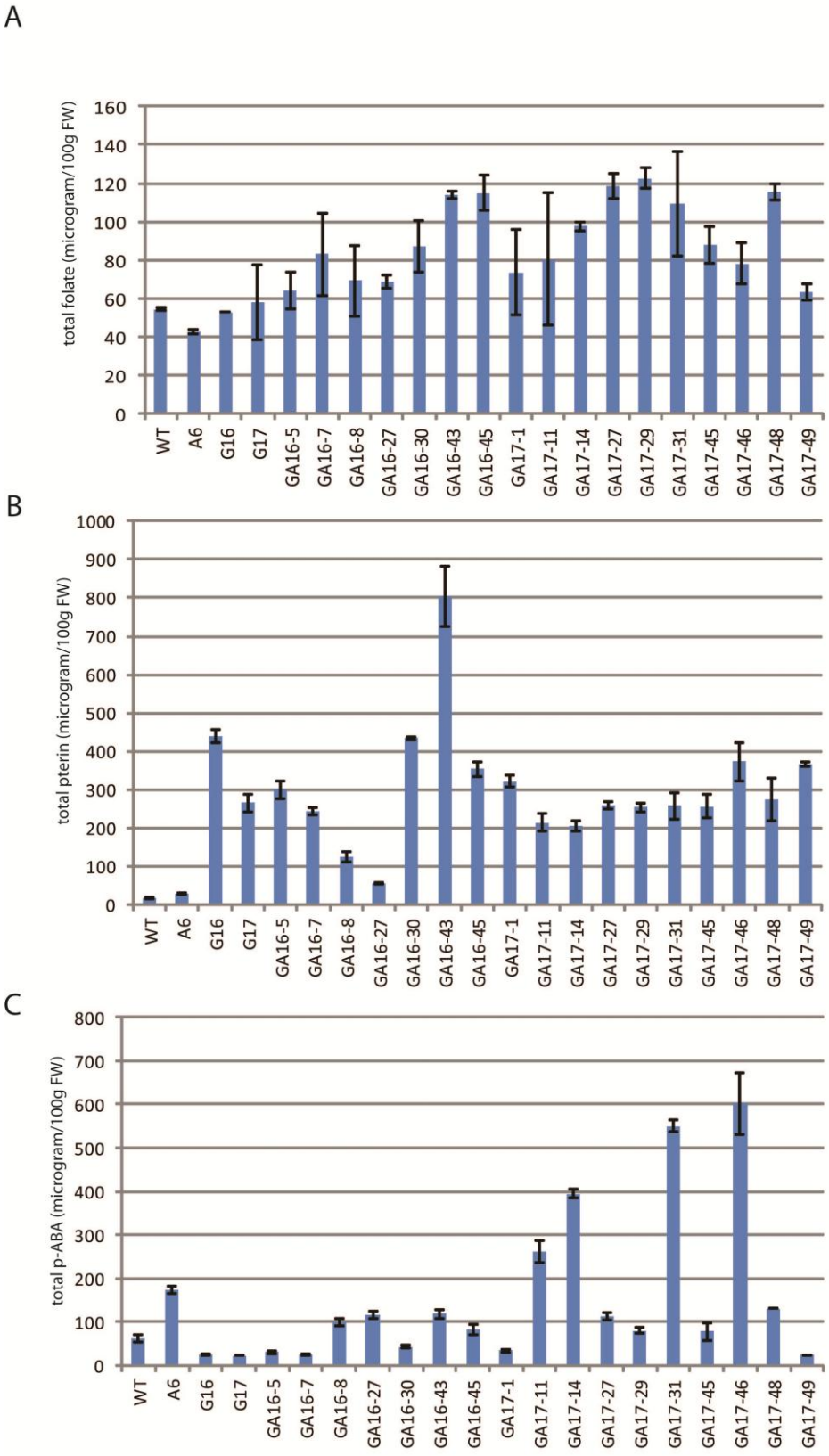


Figure 2.4: Folate (A), pterin (B) and *p*-ABA (C) content of immature transgenic potato tubers after retransformation with the pPAT_A plant transformation vector. Values are expressed in µg/100g fresh weight (FW). Values indicate averages of measurements of 2 independent samples per line. Error bars indicate standard deviation.

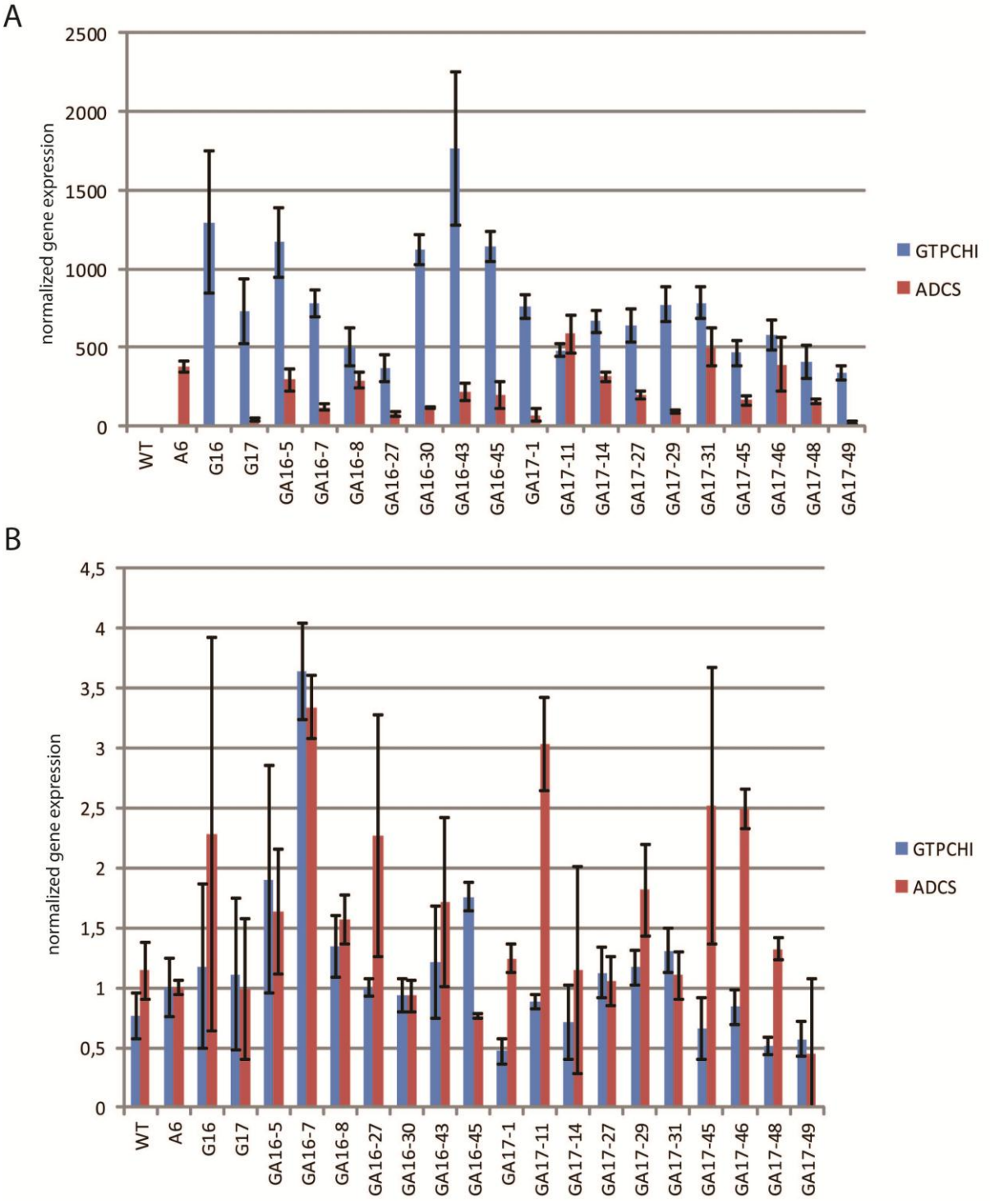


Figure 2.5 : Expression levels of *Arabidopsis thaliana* GTPCHI and ADCS transgenes (A) and the corresponding endogenous potato genes (B) in young tubers of WT and transgenic potato lines after retransformation with pPAT_A. Expression analyses were performed by real-time quantitative PCR. Potato cytoplasmic ribosomal protein L2 and elongation factor 1-α were used as reference genes for normalization. Data analysis of two independent samples was performed using qBASE, based on the $2^{-\Delta\Delta Ct}$ method. Error bars indicate standard deviation.

Feeding transgenic Arabidopsis seedlings with enhanced pterin content with p-ABA

Since a constitutive ADCS overexpression raising more than 5-fold above the wild type gene expression level could not be achieved in *Arabidopsis thaliana*, we decided to investigate whether *p*-ABA treatment of G-lines with a high pterin content could result in folate enhancement in *Arabidopsis*. Two G lines (G3 and G9, with a pterin content of 50.93 ± 66.46 and 381.11 ± 186.16 µg/100 g FW respectively) were supplemented with *p*-ABA in a *p*-ABA feeding experiment. Two feeding experiments were conducted. First, transgenic G3, G9, an empty vector control and WT *Arabidopsis* seeds were grown on half-strength MS media, supplemented with water or 100 µM of *p*-ABA. After two weeks, the seedlings were collected and *p*-ABA and folate content was analyzed (Fig. 2.6A and B). Total *p*-ABA content was strongly enhanced in all *p*-ABA supplemented samples (Fig. 2.6A), especially in the G9 line, where *p*-ABA levels up to 8396.00 ± 2256.50 µg/100 g FW were measured. Despite this massive *p*-ABA accumulation, only a slight increase in total folate content was detected in the samples grown on *p*-ABA containing media (Fig. 2.6B). In a second experiment, two week old transgenic and WT *Arabidopsis* seedlings were vacuum-infiltrated with water or 100 µM of *p*-ABA (Fig. 2.6C and D). After 4 hours of exposure, the seedlings were collected and analyzed on *p*-ABA and folate content. Total *p*-ABA content increased severely upon *p*-ABA treatment (Fig. 2.6C); again, the highest *p*-ABA levels being measured in G9 (4181.08 ± 220.61 µg/100 g FW). However, this *p*-ABA enhancement did not result in increased folate levels (Fig. 2.6D).

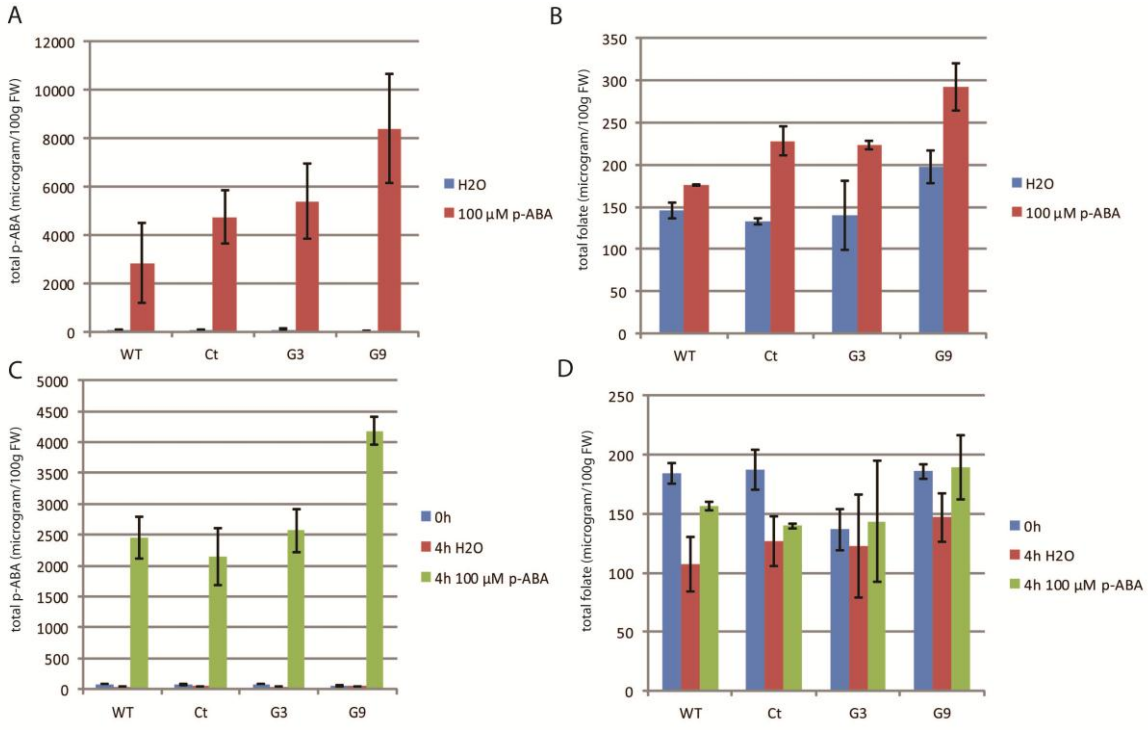


Figure 2.6: Total *p*-ABA and folate content in WT and transgenic *Arabidopsis thaliana* seedlings, supplemented with water or 100 µM of *p*-ABA. Seedlings were grown on media containing water or *p*-ABA (A and B) for two weeks. A: total *p*-ABA levels. B: total folate levels. Panel C and D show the results of 2 week old seedlings, vacuum-infiltrated for 4 hours with water or 100 µM of *p*-ABA. C: total *p*-ABA content. D: total folate content. All values are expressed in µg/100 g fresh weight (FW). Values indicate averages of measurements of 2 samples per line. Error bars indicate standard deviation. Ct: empty vector control.

Discussion

In an attempt to biofortify potato tubers with folate by metabolic engineering, *Solanum tuberosum* L. var. Desirée was stably transformed with *Arabidopsis thaliana* cDNA encoding GTPCHI and ADCS, under the control of tuber-specific patatin class 1 and granule-bound starch synthase 1 promoters respectively. This two-gene approach was previously successfully adopted in rice seeds (Storozhenko *et al.,* 2007) and in tomato fruit (Diaz de la Garza *et al.,* 2007). Although pterin content was greatly enhanced in the transgenic potato lines, folate levels were hardly doubled (Fig. 2.3A). Expression analysis revealed that this could be attributed to a lack of ADCS transgene expression (Fig. S2.1). Chimaeric sequences between the β-glucuronidase (GUS) reporter gene and the granule-bound starch synthase promoter region resulted in high GUS expression levels in potato micro tubers, which strongly exceeded the expression in potato leaves, supporting that this promoter should be highly active in potato tubers (van der Steege *et al.,* 1992). However, in some cases, upon usage of this promoter to drive specific transgene expression, it could remain inactive or at least no regenerants could be obtained with a high expression (R. Visser, unpublished results). Therefore, it was decided to retransform two G lines, enhanced in pterin content, with a new plant transformation vector to overexpress the ADCS transgene under the control of the patatin promoter, used successfully for *Arabidopsis* GTPCHI overexpression in potato tubers. Expression analysis by real-time quantitative PCR showed that in these transformants the ADCS transgene was successfully overexpressed (Fig. 2.5A and Fig. S2.2), resulting in enhanced *p*-ABA content in the corresponding tubers of some lines (Fig. 2.4C). However, folate levels in immature micro tubers again only raised modestly, up to 3-fold, as compared to the WT control (Fig. 2.4A). Since an enhancement of folate biosynthesis precursors *p*-ABA and pterins did not coincide with a similar increase of folate content, it is clear that there is another bottleneck in the folate biosynthesis pathway in potato tubers. Thus, potato differs from rice and tomato, where a similar two-gene strategy was successfully applied (Storozhenko *et al.,* 2007; Diaz de la Garza *et al.,* 2007). Additionally, it was investigated whether the overexpression of these two transgenes had an effect on the expression of the corresponding endogenous potato genes. Real-time quantitative PCR results showed that endogenous potato GTPCHI and ADCS expression patterns did not differ between WT and the transgenics (Fig. 2.5B), concluding that they were not affected by overexpression of the transgenes as previously reported in engineered tomato fruit (Waller *et al.,* 2010). Folate analysis of mature potato tubers supported the previously obtained results for the micro tubers (Fig. S2.3). Generally, folate content in mature potatoes was lower as compared to that in young tubers, which is in line with earlier reports on folate dynamics in developing potato tubers (Goyer and Navarre 2009). This study showed that folate levels decreased 2.6 to 3.4-fold between micro tubers and mature potatoes at the stage of harvest. It was proposed that this drop in folate content throughout tuber maturation is regulated at the post-transcriptional level, since the expression of the folate biosynthesis genes was sustained during tuber development (Goyer and Navarre 2009). In parallel, the same metabolic engineering strategy, combining GTPCHI and ADCS genes, was used in an attempt to enhance folate levels in whole *Arabidopsis thaliana* plants. In *Arabidopsis*, both genes were placed under the control of the constitutive 35S promoter. The GTPCHI transgene was successfully overexpressed, resulting in increased pterin content in the *Arabidopsis* transgenics. However, as observed in the first transformation round in potato (Fig. S2.1), overexpression of the ADCS transgene could not be detected (Fig. S2.4 and S2.5) and therefore did not lead to an enhancement of folate levels (Fig. S2.6). Thus, it was decided to conduct *p*-ABA feeding experiments on two *Arabidopsis* G lines to investigate whether pterin and *p*-ABA enhancement could

possibly result in an increase in folate content. Only a modest folate increase was detected upon chronic treatment of *Arabidopsis* seedlings with 100 µM of *p*-ABA (Fig. 2.6B). Altogether, the data suggest that enhancement of pterin and *p*-ABA content in potato tubers and in *Arabidopsis* is not sufficient to successfully biofortify these species with folate, in contrast with the results obtained in rice (Storozhenko *et al*., 2007) and tomato (Diaz de la Garza *et al*., 2007), and point to the presence of other mechanisms regulating folate status in potato and *Arabidopsis*, which are different, lacking or of minor importance in rice seeds and tomato fruit. Unfortunately, little is known about the ability and tools that plants use to regulate folate levels (Blancquaert *et al.,* 2010; Hanson and Gregory 2011). Most likely, folate enhancement in potato tubers and *Arabidopsis* plants is hampered due to the presence of one or more limiting steps in the folate biosynthesis pathway, downstream pterin and *p*-ABA biosynthesis. Since the production of tetrahydrofolate occurs in the mitochondria and requires the import of its precursors into the latter, it is possible that folate enhancement is restricted due to restricted import of *p*-ABA and pterin into this organelle. However, it was reported that the import of *p*-ABA into the mitochondria occurs by simple diffusion (Quinlivan *et al*., 2003). On the other hand, although little is known about pterin transport mechanisms in plants, it is most likely mediated by specific pterin carriers. The existence of such a pterin transporter has been reported in *Leishmania* (BT1) and it appears to be a member of the folate-biopterin transporter (FBT) family (Lemley *et al*., 1999). The *Arabidopsis* genome possesses nine family members, of which one was already characterized as being important in folate transport (Lemley *et al*., 1999). It is possible that another member of this *Arabidopsis* FBT family encodes a specific pterin carrier (Hanson and Gregory 2011). Therefore, characterization of this pterin transporter in plants may provide a clue as to whether the pterin transporter is indeed a rate-limiting step in biofortifying crops with folates. Besides insufficient import of *p*-ABA and pterins into the mitochondria, it is possible that the bottleneck hampering folate accumulation in *Arabidopsis* and potato tubers is a limitation in activity of one or more tetrahydrofolate biosynthesis enzymes present in the mitochondria. Overexpression of dihydropterin pyrophosphokinase – dihydropteroate synthase (HPPK-DHPS), which couples *p*-ABA to the activated pterin form 6-hydroxymethyldihydropterin pyrophosphate (HMDHP-P_2) could push the flux of precursors toward folates, since it is known to be a regulatory point in folate biosynthesis (Mouillon *et al.,* 2002). In this respect, it was reported that the DHPS domain of this bifunctional plant enzyme is feedback inhibited by folate biosynthesis intermediates dihydropteroate, dihydrofolate monoglutamate and tetrahydrofolate monoglutamate (Mouillon *et al*., 2002). On the other hand, overexpression of folylpolyglutamate synthetase (FPGS), which performs the last step in tetrahydrofolate biosynthesis, could possibly 'pull' the folate biosynthesis pathway toward a folate accumulation in *Arabidopsis* and potato and could avoid any feedback inhibition of folate intermediates on its biosynthesis enzymes. Indeed, expression of FPGS in folate engineered tomato fruit was up-regulated (Waller *et al*., 2010), which could explain why the two-gene overexpression strategy was sufficient to enhance folate levels. Since folate-dependent enzymes prefer polyglutamylated folates above their monoglutamate forms (Shane 1989) and since polyglutamylation favors cellular retention (Appling 1991), overexpression of FPGS, in combination with GTPCHI and ADCS, could have additional advantages, such as increasing folate stability. However, although this three-gene strategy could result in a successful folate enhancement, a higher fraction of polyglutamylated folates cannot be guaranteed, due to the existence of gamma-glutamyl hydrolase (GGH) activity in plants (Akhtar *et al*., 2010), which cleaves the polyglutamate tail of polyglutamylated folates, reconverting them to their monoglutamate forms. The three-gene strategy could thus be more successful in biofortifying plants with folates in a background with decreased

GGH activity. Finally, besides an insufficient capacity of the folate biosynthesis enzymes downstream the *p*-ABA and pterin branch, it is possible that their activity is controlled by a yet unknown regulatory mechanism in potato tubers and *Arabidopsis* plants.

Overall, we conclude that metabolic engineering of the *p*-ABA and pterin branch alone is not sufficient to successfully biofortify potato tubers and *Arabidopsis* plants with folate due to the presence of another constraint downstream the folate biosynthesis pathway. This bottleneck is lacking in rice seeds and tomato fruit and possibly explains why half a decade after successful biofortification of rice and tomato, still no other crops have been reported to be engineered to contain substantially higher folate levels, besides a few single-gene (GTPCHI) overexpressing plants with a modest folate increase (Nunes *et al.,* 2009; Naqvi *et al.,* 2009). It is clear that the regulation of folate biosynthesis is quite complex (reviewed in Ravanel et al., 2011) and a better understanding of this pathway is required in order to determine an engineering strategy that could be generalized to most staple crops. However, the hurdles and possible solutions reported here may help to design strategies to biofortify other staple crops.

Acknowledgments

The authors thank Marjan Bergervoet (Wageningen UR Plant Breeding, The Netherlands) for her expert assistance in potato transformation.
DVDS and WL gratefully acknowledge financial support from Ghent University (Bijzonder Onderzoeksfonds, BOF2004/GOA/012 and BOF2009/G0A/004), and the Fund for Scientific Research Flanders (FWO, project 3G012609).
D.B. is indebted to FWO (Fonds voor Wetenschappelijk Onderzoek – Vlaanderen), for a PhD grant.

Supplementary data

Table S2.1: List of primers used for cloning and real-time quantitative PCR

Primer name	Primer sequence (5'->3')
STOSER 136	TTAACCATGGGCGCATTAGATGAGGGA
STOSER 137	TAGCCATGGTCAAAATGGAGAGCTTGACTC
STOSER 138	TATGGATCCCTAAACGAGTTATGAACATGAA
STOSER 139	TACGGATCCAAAACTATTGTCTCCTCTGAT
STOSER 178	CCGAGCTCATTGTGTTACATATTTTACTTTTGACT
STOSER 179	CTACTAGTGTGTTTTTAAATTTTGTTGGTGCTTTG
STOSER 28	TTTGTGAGGACTTTGTTGATTGATAA
STOSER 29	GAGAAATACCCGATTGAGCCAGAGT
STOSER 42	AAAAAGCAGGCTCTACCATGGGCGCATTAGATGAGGGA
STOSER 43	AGAAAGCTGGGTCTTAGTTCTTTGAACTAGTGTTTCGCTG
STOSER 39	TGTCTTTATGCTCTTCGTCCCCTGA
STOSER 73	AGAAAGCTGGGTAAAACTATTGTCTCCTCTGATCACT
STOSER 150	AAGAATGTTGACAAGAAG
STOSER 151	CAGAGAACTAGAACCTAA
STOSER 152	GCTATGATTGGTCAGGTT
STOSER 153	CACACGGTACTTATGGTAA
STOSER 154	ATTATCACCTGAAGTTGAC
STOSER 155	GTGTAGCAATGAGTTCTT
STOSER 156	TGATTGTTGACCTTCTAA
STOSER 157	ACTGTTGTGTATGATTCT
pot GTPCHI 1 F	TGAGTCTTGCTTGCTTCC
pot GTPCHI 1 R	ACCTGTTGGCTTGATTCC
pot ADCS 1F	TGGATAGGAATGCTATGCTGGAAG
pot ADCS 1R	GCTCACATACACGCCCAAGG
ACTIN11 F	CCACATGCTATTCTGCGTTTGGACC
ACTIN11 R	CATCCCTTACGATTTCACGCTCTGC
GapC F	GAGGGTGGTGCAAGAAGGTT
GapC R	AGGGGAGCAAGGCAGTTAGTGC

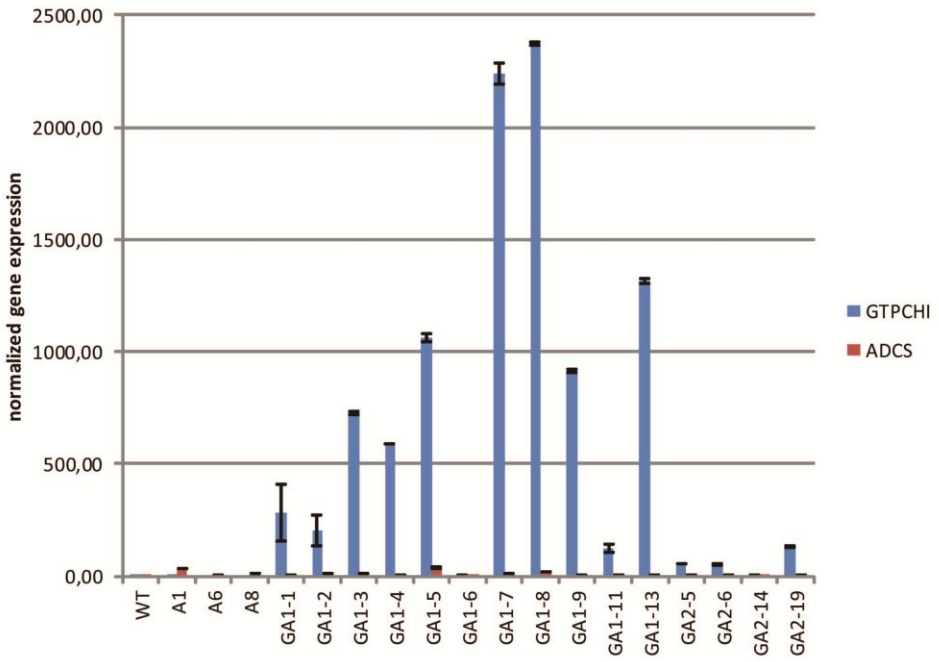


Figure S2.1: Expression levels of *Arabidopsis thaliana* GTPCHI and ADCS transgenes in WT and transgenic potato tubers transformed with pGBSSI_A (A lines), pGBSSI_A_PAT_G1 (GA1) and pGBSSI_A_PAT_G2 (GA2).

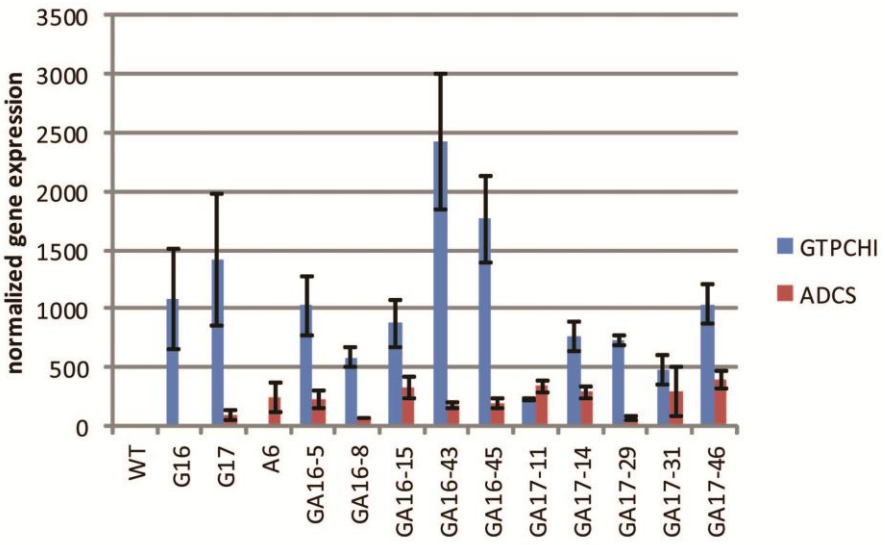


Figure S2.2: Expression levels of *Arabidopsis thaliana* GTPCHI and ADCS transgenes in WT and transgenic mature potato tubers after retransformation with the pPAT_A plant transformation vector.

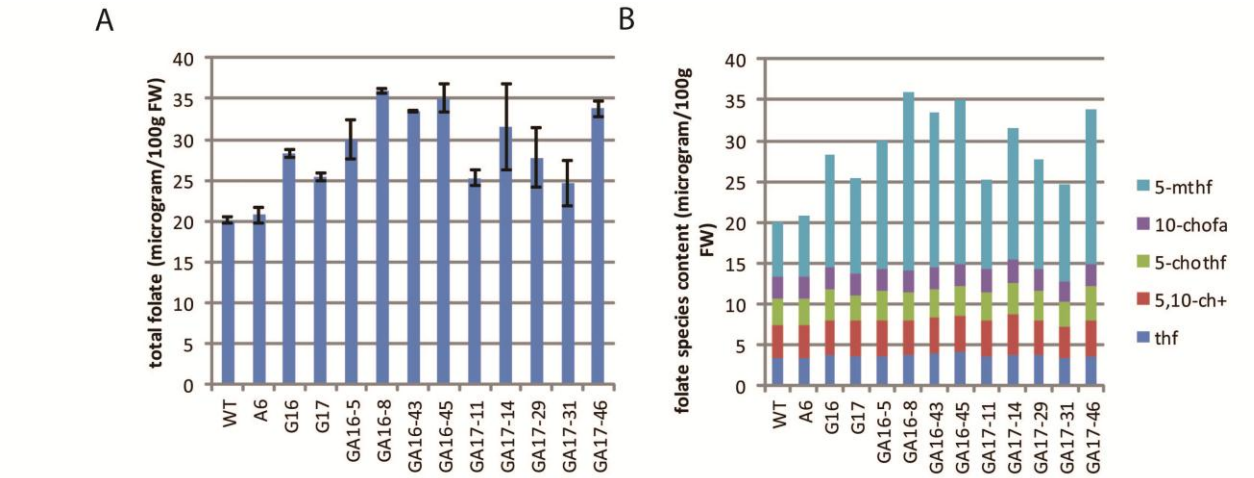


Figure S2.3: Total folate content (A) and distribution of the different folate forms (B) in mature potatoes after retransformation with pPAT_A.

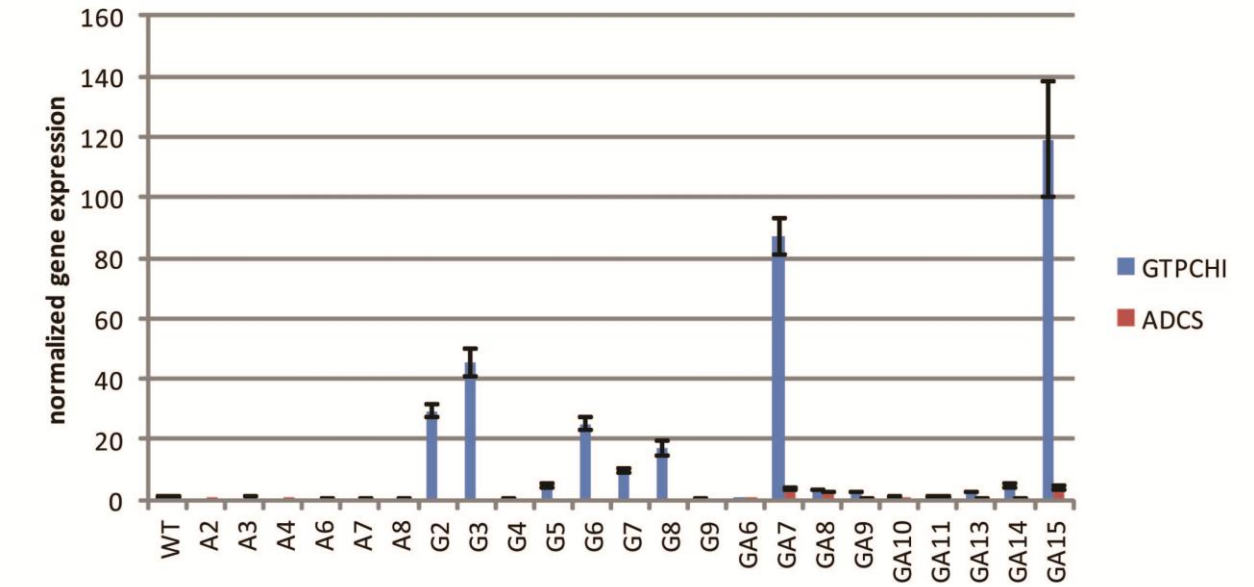


Figure S2.4: : Expression levels of GTPCHI and ADCS in WT and transgenic *Arabidopsis thaliana* leaves after transformation with the pATH_A (A lines), pATH_G (G lines) and pATH_GA (GA lines) plant transformation vectors.

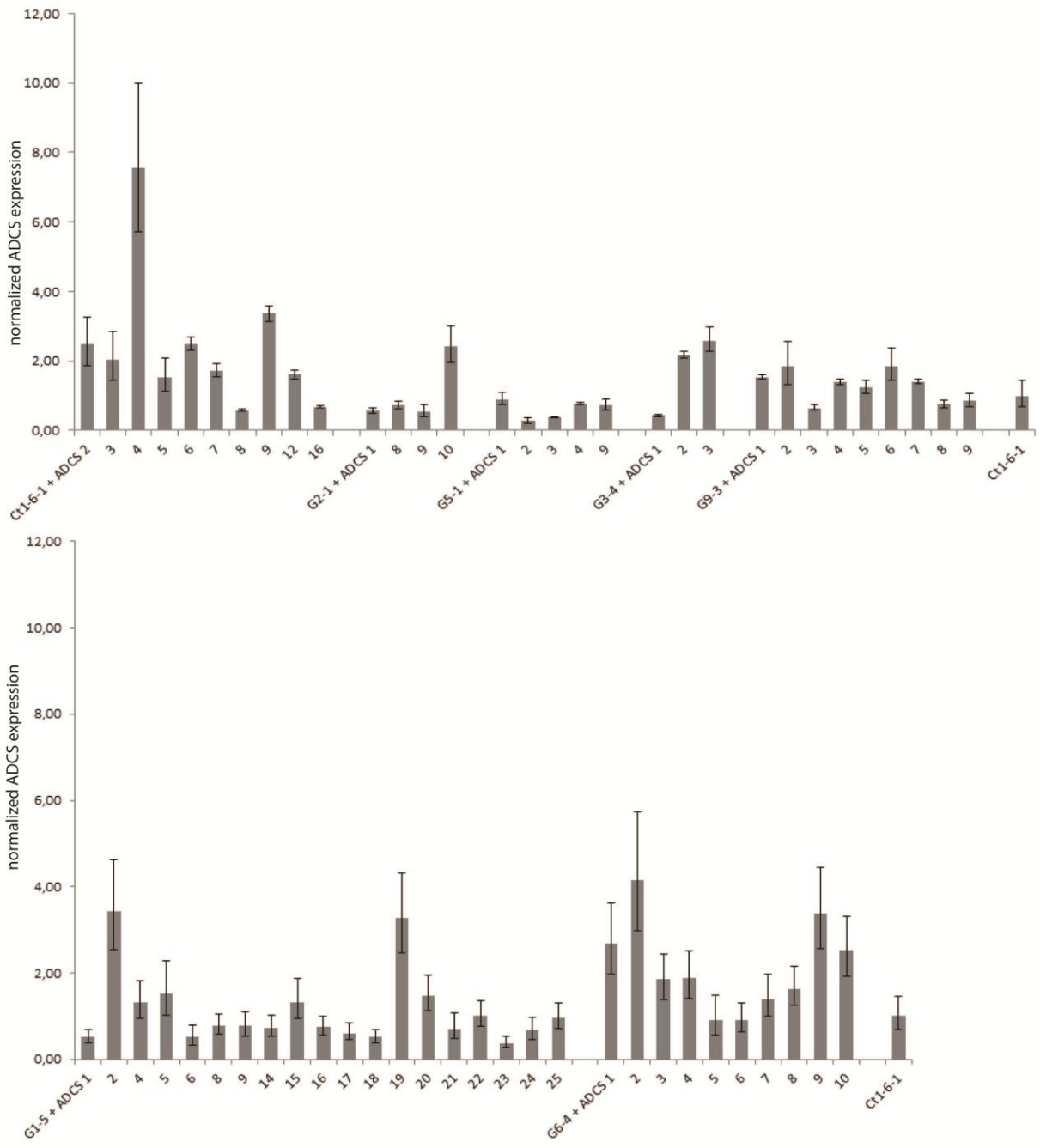


Figure S2.5: ADCS expression levels in empty vector control and transgenic *Arabidopsis thaliana* leaves after retransformation of 6 G lines and an empty vector control line with the pADCSH7 plant transformation vector.

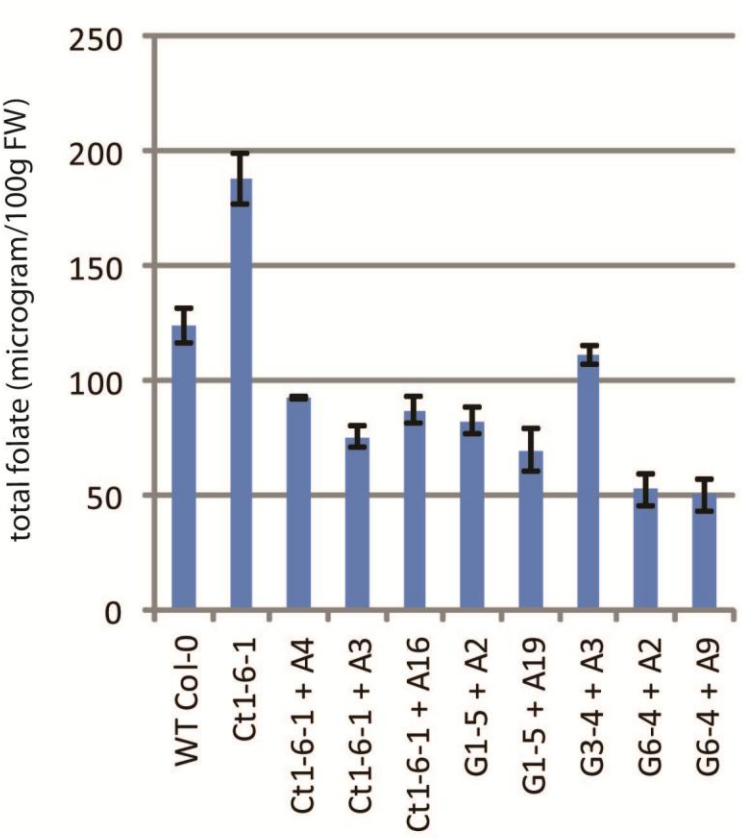


Figure S2.6: Total folate content in transgenic G+A, A, empty vector control (Ct1-6-1) and WT *Arabidopsis thaliana* leaves, after retransformation with the pADCSH7 plant transformation vector.

Supplementary material and methods

p-ABA, pterin and folate analysis

All chemicals, materials, stock and working solutions used and prepared for folate, *p*-ABA and pterin analyses were as previously described (De Brouwer *et al.,* 2010; Navarrete *et al.,* 2012). Preparation of *Arabidopsis* samples was done according to Navarrete *et al.,* 2012.

1. Preparation of potato samples

During sample preparation prior to folate and pterin analyses, all manipulations were carried out under subdued light. For folate determination, 1.5 ml of extraction buffer (containing the internal standards) was added to ± 500 mg of potato flesh and the capped Eppendorf® tube was incubated at 100°C for 10 minutes before homogenization in a Retsch® Mill (Haan, Germany). A trienzyme treatment with 30 µl α-amylase (23.5 units/ µl) (20 minutes), 150 µl protease (≥ 3.5 units/ mg) (1 hour at 37 °C) and 100 µl conjugase (2 hours at 37°C) was used to degrade the starchy matrix, release protein-bound folates and deconjugate polyglutamylated folates. To stop protease and conjugase activity, additional heat treatments (10 minutes at 100°C) were carried out, followed by cooling on ice. The resulting solution was ultra-filtrated at 12851 x *g* for 15 min. For pterin analysis, a similar procedure was followed. In this case, a dienzyme treatment was applied, with 30 µl amylase (1 hour at 37°C) and 150 µl protease (1 hour at 37°C) followed by a heat treatment (10 minutes at 100°C) and ultra-filtration.

For *p*-ABA analysis, ± 500 mg of potato flesh was ground after boiling in 200 µl of water. The samples were then transferred to 15-ml Falcon® tubes and diluted with a 11.8 ml volume of a methanolic extraction solvent containing 3-amino-4-methylbenzoic acid as internal standard. After extraction, the tubes were centrifuged at 2655 x g for 15 min. The supernatant was collected and 5 ml of the methanolic layer was dried completely under a nitrogen gas flow at 45°C. One ml of water was added to the residue, followed by sonication for 5 min. To a volume of 0.4 ml of this reconstituted residue, 50 µl of 2 M HCl was added, after which the samples were incubated at 80°C for 2 hours. After cooling down, 50 µl of 2 M NaOH was added for neutralization. Finally, all sample solutions were ultra-filtrated.

2. UPLC™ conditions

The chromatography materials and conditions were as previously described (De Brouwer *et al.,* 2010; Navarrete *et al.,* 2012), with minor modifications:

For folate analysis (column: 60°C), the mobile phase consisted of eluent A (0.1% formic acid in water) and eluent B (0.1% formic acid in acetonitrile) and was pumped at a flow rate of 0.6 ml/min. The starting condition (100% A) was kept for 1 minute. Subsequently, there was a nonlinear increase to 95% in 3 minutes, where it was kept for 1 minute. The mobile phase was then immediately adjusted to its initial composition and held for 3 minutes for re-equilibration, resulting in a total time of 8 minutes.

For pterin analysis (column: 30°C), the mobile phase consisted of eluent A (0.1% acetic acid in water) and eluent B (0.1% acetic acid in acetonitrile) and was pumped at a flow rate of 0.45 ml/min. The starting condition (100% A) was decreased linearly to 85% in 2.5 minutes, followed by a nonlinear decrease to 0% in 1.5 minutes, where it was kept for 1.0 minute. Subsequently, the mobile phase was adjusted to its initial composition, again in a nonlinear way, and held for 2 minutes for re-equilibration, resulting in a total time of 7 minutes.

For *p*-ABA analysis, (column: 40°C), the mobile phase consisted of eluent A (0.1% formic acid in water) and eluent B (0.1% formic acid in acetonitrile) and was pumped at a flow rate of 0.6 ml/min. The starting condition (100% A) was immediately decreased in a nonlinear way to 30% in 2 minutes, followed by a decrease to 0% over 1 minute, where it was kept for 2 minutes. Subsequently, the mobile phase was adjusted to its initial composition and held for 2 minutes for re-equilibration, resulting in a total time of 7 minutes.

3. Mass spectrometric instrumentation and settings

The detection was performed by ESI utilizing heated auxiliary gas (Turbospray), in the *scheduled* MRM™ mode, using an Applied Biosystems API 4000 tandem mass spectrometer (Ontario, Canada). For folate and pterin analysis, the source conditions were as follows: source temperature at 500°C, ionspray voltage at 3.5 kV and -4.5 kV, respectively. For p-ABA analysis the settings were as follows: source temperature at 550°C, ionspray voltage at 5.5 kV. In all instances, Q1 and Q3 were at unit mass resolution. The interface heater was on in all cases (± 100°C); nitrogen was used for the nebulizer (gas 1), heater (gas 2), curtain and collisionally activated dissociation (CAD) gas. Gasses 1, 2 and curtain gas had respective pressure settings at 70, 90 and 25 psig for folate analysis, 80, 90 and 26 psig for pterin analysis, and 75, 20 and 30 psig for p-ABA analysis. Vacuum settings for CAD for analysis of folates, pterins and p-ABA were at 8, 8 and 6, respectively.

References

Akthar TA, Orsomando G, Mehrshahi P, Lara-Nunez A, Bennett MJ, Gregory JFIII, Hanson AD. 2010. A central role for γ-glutamyl hydrolases in plant folate homeostasis. The Plant Journal **64**, 256-266.

Appling DR. 1991. Compartmentation of folate-mediated one-carbon metabolism in eukaryotes. The Journal of the Federation of American Societies for Experimental Biology **5**, 2645-2651.

Asensi-Fabado MA, Munné-Bosch S. 2010. Vitamins in plants: occurrence, biosynthesis and antioxidant function. Trends in Plant Science **10**, 582-592.

Bekaert S, Storozhenko S, Mehrshahi P, Bennett MJ, Lambert W, Gregory JFIII, Schubert K, Hugenholtz J, Van Der Straeten D, Hanson AD. 2008. Folate biofortification in food plants. Trends in Plant Science **13**, 28-34.

Blancquaert D, Storozhenko S, Loizeau K, De Steur H, De Brouwer V, Viaene J, Ravanel S, Rébeillé F, Lambert W, Van Der Straeten D. 2010. Folates and folic acid: from fundamental research toward sustainable health. Critical Reviews in Plant Science **29**, 14-35.

Burdge GC, Lillycrop KA. 2012. Folic acid supplementation in pregnancy: are there devils in the detail? British Journal of Nutrition **108**, 1924-1930.

Choi SW, Friso S. 2005. Interactions between folate and aging for carcinogenesis. Clinical Chemistry and Laboratory Medicine **43**, 1151-1157.

Clough SJ, Bent AF. 1998. Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. The Plant Journal **16**, 735-743.

Cook RJ. 2001. Disruption of histidine catabolism in NEUT2 mice. Archives of Biochemistry and Biophysics **392**, 226-232.

De Brouwer V, Storozhenko S, Stove CP, Van Daele J, Van Der Straeten D, Lambert W. 2010. Ultra-performance liquid chromatography-tandem mass spectrometry (UPLC-MS/MS) for the sensitive determination of folates in rice. Journal of Chromatography B – Analytical Technologies in the Biomedical and Life Sciences **878**, 509-513.

De Steur H, Gellynck X, Storozhenko S, Ge LQ, Lambert W, Van Der Straeten D, Viaene J. 2010. Health impact in China of folate-biofortified rice. Nature Biotechnology **28**, 554-556.

Diaz de la Garza R, Quinlivan EP, Klaus SM, Basset GJ, Gregory JFIII, Hanson AD. 2004. Folate biofortification in tomatoes by engineering the pteridines branch of folate synthesis. Proceedings of the National Academy of Sciences of the United States of America **101**, 13720-13725.

Diaz de la Garza RI, Gregory JFIII, Hanson AD. 2007. Folate biofortification of tomato fruit. Proceedings of the National Academy of Sciences of the United States of America **104**, 4218-4222.

Endres M, Ahmadi M, Kruman I, Biniszkiewicz D, Meisel A, Gertz K. 2005. Folate deficiency increases postischemic brain injury. Stroke **36**, 321-325.

Filati J, Kiser J, Rose R, Comai L. 1987. Efficient transfer of a glyphosphate tolerance gene into tomato using a binary *Agrobacterium tumefaciens* vector. Biotechnology **5**, 726–730.

Geisel J. 2003. Folic acid and neural tube defects in pregnancy – a review. The Journal of Perinatal and Neonatal Nursing **17**, 268-279.

Goyer A, Navarre DA. 2009. Folate is higher in developmentally younger potato tubers. Journal of the Science of Food and Agriculture **89**, 579-583.

Goyer A, Sweek K. 2011. Genetic diversity of thiamin and folate in primitive cultivated and wild potato (*Solanum*) species. Journal of Agricultural and Food Chemistry **59**, 13072-13080.

Hanson AD, Roje S. 2001. One-carbon metabolism in higher plants. Annual Review of Plant Physiology and Plant Molecular Biology **52**, 119-137.

Hanson AD, Gregory JFIII. 2011. Folate biosynthesis, turnover, and transport in plants. Annual Review of Plant Biology **62**, 105-125.

Hellemans J, Mortier G, De Paepe A, Speleman F, Vandesompele J. 2007. qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. Genome Biology **8**, R19.

Hoekema A, Huisman MJ, Molendijk L, Elzen PJM, Cornelissen BJC. 1989. The genetic engineering of two commercial potato cultivars for resistance to potato virus X. Nature Biotechnology **7**,273–278.

Hossain T, Rosenberg I, Selhub J, Kishore G, Beachy R, Schubert K. 2004. Enhancement of folate in plants through metabolic engineering. Proceedings of the National Academy of Sciences of the United States of America **101**, 5158-5163.

Hubner R, Houlston RD, Muir KR. 2007. Should folic acid fortification be mandatory? No British Medical Journal **334**, 1253.

Jiang L, Liu Y, Sun H, Han Y, Li J, Li C, Guo W, Meng H, Li S, Fan Y, Zhang C. 2013. The mitochondrial folylpolyglutamate synthetase gene is required for nitrogen utilization during early seedling development in *Arabidopsis*. Plant Physiology **161**, 971-989.

Karimi M, Inzé D, Depicker A. 2002. GATEWAY vectors for *Agrobacterium*-mediated plant transformation. Trends in Plant Science **7**, 193-195.

Lemley C, Yan S, Dole VS, Madhubala R, Cunningham ML, Beverley SM, Myler PJ, Stuart KD. 1999. The Leishmania donovani LD1 locus gene ORFG encodes a biopterin transporter (BT1). Molecular and Biochemical Parasitology **104**, 93-105.

Li GM, Presnell SR, Gu, LY. 2003. Folate deficiency, mismatch repair-dependent apoptosis, and human disease. Journal of Nutrional Biochemistry **14**, 568-575.

Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta Ct}$ method. Methods **25**, 402-408.

Mouillon JM, Ravanel S, Douce R, Rébeillé F. 2002. Folate synthesis in higher plant mitochondria : coupling between the dihydropterin pyrophosphokinase and the dihydropteroate synthase activities. Biochemistry Journal **363**, 313-319.

Naqvi S, Zhu C, Farre G, Ramessar K, Bassie L, Breitenbach J, Perez Conesa D, Ros G, Sandmann G, Capell T, Christou P. 2009. Transgenic multivitamin corn through biofortification of endosperm with three vitamins representing three distinct metabolic pathways. Proceedings of the National Academy of Sciences of the United States of America **106**, 7762-7767.

Naumkina E, Bolyakina Y, Romanov G. 2007. Organ-specific and inducibility of patatin class I promoter from potato in transgenic *Arabidopsis* plants. Russian Journal of Plant Physiology **54**, 350-359.

Navarrete O, Van Daele J, Stove C, Lambert W, Van Der Straeten D, Storozhenko S. 2012. A folate independent role for cytosolic HPPK/DHPS upon stress in *Arabidopsis thaliana*. Phytochemistry **73**, 23-33.

Nicot N, Hausman JF, Hoffmann L, Evers D. 2005. Housekeeping gene selection for real-time RT-PCR normalization in potato during biotic and abiotic stress. Journal of Experimental Botany **56**, 2907-2914.

Nunes ACS, Kalkmann DC, Aragão FJL. 2009. Folate biofortification of lettuce by expression of a codon optimized chicken GTP cyclohydrolase I gene. Transgenic Research **18**, 661-667.

Papakostas G, Cassiello CF, Iovieno N. 2012. Folates and S-adenosylmethionine for Major Depressive Disorder. Canadian Journal of Psychiatry **57**, 406-413.

Pel MA, Foster SJ, Park TH, Rietman H, van Arkel G, Jones JDG, Van Eck HJ, Jacobsen E, Visser RGF, Van der Vossen EAG. 2009. Mapping and cloning of late blight resistance genes from *Solanum venturii* using an interspecific candidate gene approach. Molecular Plant-Microbe Interactions **22**, 601–615.

Quinlivan EP, Roje S, Basset G, Shachar-Hill Y, Gregory JFIII, Hanson AD. 2003. The folate precursor p-aminobenzoate is reversibly converted to its glucose ester in the plant cytosol. Journal of Biological Chemistry **278**, 20731-20737.

Ravanel S, Douce R, Rébeillé F. 2011. Metabolism of folates in plants. Advances in Botanical Research **59**, 67-106.

Romanov GA, Naumkina EM, Ashapkin VV, Vanyushin F. 2007. Methylation of GCGG sites of the patatin promoter is organ-specific and inversely correlates with its activity. Doklady Biochemistry and Biophysics **417**, 327-330.

Scott J, Rébeillé F, Fletcher J. 2000. Folic acid and folates: the feasibility for nutritional enhancement in plant foods. Journal of the Science of Food and Agriculture **80**, 795 – 824.

Scott JM, Weir DG. 1996. Homocysteine and cardiovascular disease. Quarterly Journal of Medicine **89**, 561-563.

Seshadri S, Beiser A, Selhub J, Jacques PF, Rosenberg IH, D'Agostino RB, Wilson PWF, Wolf PA. 2002. Plasma homocysteine as a risk factor for dementia and Alzheimer's disease. The New England Journal of Medicine **346**, 476-483.

Shane B. 1989. Folylpolyglutamate synthesis and role in the regulation of one-carbon metabolism. Vitamins and Hormones **45**, 263-335.

Stanger O. 2004. The potential role of homocysteine in percutaneous coronary interventions (PCI): review of current evidence and plausibility of action. Cellular and Molecular Biology **50**, 953-988.

Storozhenko S, De Brouwer V, Volckaert M, Navarrete O, Blancquaert D, Zhang GF, Lambert W, Van Der Straeten D. 2007. Folate fortification of rice by metabolic engineering. Nature Biotechnology **25**, 1277-1279.

van der Steege G, Nieboer M, Swaving J, Tempelaar MJ. 1992. Potato granule-bound starch synthase promoter-controlled GUS expression: regulation of expression after transient and stable transformation. Plant Molecular Biology **20**, 19-30.

Waller JC, Akhtar TA, Lara-Nunez A, Gregory JFIII, McQuinn RP, Giovannoni JJ, Hanson AD. 2010. Developmental and feedforward control of the expression of folate biosynthesis genes in tomato fruit. Molecular Plant **3**, 66-67.

Chapter 3

Increasing folate stability in biofortified rice by enhancing folate polyglutamylation and through complexation with folate binding proteins

Vector cloning and molecular analyses were performed by Dieter Blancquaert. Cloning of the rice glutelin B4 promoter and construction of vectors GAmtF and GAsFBP was done by Sergei Storozhenko. Folate analyses were performed by Jeroen Van Daele and Filip Kiekens (Faculty of Pharmaceutical Sciences, Laboratory of Toxicology).

Increasing folate stability in biofortified rice by enhancing folate polyglutamylation and through complexation with folate binding proteins

Dieter Blancquaert[a], Jeroen Van Daele[b], Sergei Storozhenko[a], Filip Kiekens[b], Christophe Stove[b], Willy Lambert[b], Dominique Van Der Straeten[a,#]

[a] Laboratory of Functional Plant Biology, Department of Physiology, Ghent University, K.L. Ledeganckstraat 35, 9000 Gent, Belgium

[b] Laboratory of Toxicology, Ghent University, Harelbekestraat 72, 9000 Gent, Belgium

ABSTRACT

Improving nutritional health is one of the major socio-economic challenges of the 21st century, particularly in view of the continuously growing and aging world population. Folate deficiency is an important and underestimated problem of micronutrient malnutrition affecting billions of people worldwide. More and more countries are adapting policies to fight folate deficiency, mostly by fortifying foods with folic acid. However, there is growing concern about this practice, calling for alternative or complementary strategies. Enhancing folate content in staple crops by metabolic engineering is a promising, highly cost-effective strategy to eradicate folate malnutrition worldwide. Over the last decade, major progress has been made in this field, with a proof of concept in tomato and rice, most effective upon the combination of two genes in folate biosynthesis. Unfortunately, folates are unstable molecules, which degrade easily upon storage. Therefore, folate enhancement could loose its efficiency, especially in staple crops which are subjected to a long-term storage, such as rice. Indeed, we observed a significant reduction of high folate levels upon storage of biofortified rice at different temperatures and designed two strategies to counteract this degradation. A first approach comprised the enhancement of folate polyglutamylation in biofortified rice by endosperm-specific overexpression of mitochondrial or cytosolic folylpolyglutamate synthase (FPGS). In a second approach, we tried to increase folate stability through complexation of folates with folate binding proteins (FBP).

Introduction

Folate is a water-soluble vitamin (B9), which can only be synthesized *de novo* by plants and microorganisms. Thus, humans are entirely dependent on their diet to obtain the necessary dose of folates needed for a broad range of physiological and molecular processes. The recommended daily allowance (RDA) of folates is 400 µg for adults and 600 µg for pregnant women (National Institutes of Health – Office of Dietary Supplements). Folates act as one-carbon (C1) donors and acceptors, in a series of reactions called C1 metabolism. In addition, folates play a central role in the biosynthesis and metabolism of nucleotides, amino acids (serine, glycine, histidine and methionine) and pantothenate (vitamin B5) (for a review, see Blancquaert *et al.*, 2010) and provide methyl units to methyltransferases, which use a broad range of substrates, such as hormones, DNA, proteins and lipids, as part of the methyl cycle (Scott *et al.*, 2000). Green leafy vegetables, beans and certain fruits are rich sources of folates, as are fermented products. However, most staple crops, although rich in starch content, contain a low folate level while populations consuming monotonous diets, mainly consisting of these staple crops, often suffer from a suboptimal folate intake. Folate deficiency can cause neural tube defects (NTDs, such as *spina bifida* and *anencephaly*) (Geisel, 2003) and megaloblastic anaemia (Li *et al.*, 2003) and is associated with a higher risk on cardiovascular (Scott and Weir, 1996) and coronary diseases (Stanger, 2004), Alzheimer disease (Seshadri *et al.*, 2002), stroke (Endres *et al.*, 2005), several cancers (Choi and Friso, 2005) and major depressive disorder (Papakostas *et al.*, 2012). Several strategies are currently available to reduce folate deficiency, but they require educational efforts, changes in dietary habits and/or specialized infrastructure, making them difficult to implement in poorer regions of the world. Folate biofortification of crops through metabolic engineering can offer a sustainable alternative to fight folate deficiency, especially for poor populations in rural remote areas. Thus far, metabolic engineering was applied solely through the overexpression of key folate biosynthesis genes. Over the past decade, engineering attempts were reported in *Arabidopsis* (Hossain *et al.*, 2004; Blancquaert *et al.*, 2013a), tomato (Diaz de la Garza *et al.*, 2004; Diaz de la Garza *et al.*, 2007), rice (Storozhenko *et al.*, 2007), lettuce (Nunes *et al.*, 2009), white corn (Naqvi *et al.*, 2009) and potato (Blancquaert *et al.*, 2013a). Co-overexpression of both GTP cyclohydrolase I (GTPCHI) and aminodeoxychorismate synthase (ADCS) proved to be the most successful approach thus far, with a proof of concept in tomato fruit (Diaz de la Garza *et al.*, 2007) and rice seeds (Storozhenko *et al.*, 2007), where an increase in folate content up to 25 (tomato) and 100 fold (rice) could be detected.

Another important issue with respect to folate biofortification is folate stability, since obtaining high levels of this vitamin would be meaningless if they drop to basal levels upon food storage and processing. Indeed, folates are unstable compounds, susceptible to oxidative and photo-oxidative catabolism (Scott *et al.*, 2000) and degradation by pH variations (most folates are stable at pH values between 4 and 8 at 37°C, except THF and dihydrofolate) (De Brouwer *et al.*, 2007)). Therefore, these factors need to be taken into account when considering engineered staple crops as a way to fight folate deficiency. Since polyglutamylation enhances the anionic nature of the vitamin, hence improving cellular retention, increasing the ratio between folate polyglutamates and monoglutamates could enhance its stability. Moreover, folate dependent enzymes have a preference for binding with folate polyglutamylated forms and protein activity is positively correlated with the length of the glutamate tail (Shane, 1989). In WT rice seeds half of the folate pool is polyglutamylated, whereas only 10 % is in

engineered rice (Storozhenko *et al.*, 2007). In this study, we tried to enhance folate polyglutamylation in biofortified rice, by overexpression of mitochondrial or cytosolic FPGS. In a second approach to enhance folate stability, we attempted to promote complexation of folates with folate binding proteins (FBP) through an endosperm specific overexpression of a synthetic, soluble FBP (sFBP). Finally, these two aforementioned strategies were combined. Although conclusions on folate stability in these newly created rice transgenics remain inconclusive, we observed that folate levels could be further enhanced in biofortified rice seeds, as compared to those in the previously reported GA (GTPCHI + ADCS) approach (Storozhenko *et al*., 2007).

Materials and Methods

Molecular cloning and construct design

Cloning of the rice glutelin B4 promoter and cDNA

The expression of the three different FBP transgenes (sFBP, CAFBP and GluB4FBP, see below) was controlled by the glutelin B4 promoter. In order to obtain this promoter, the full rice glutelin B4 gene (AACV01003830) was cloned by amplification with primers STOSER 121 and 122 (all primers are listed in Table 3.1) (Supplemental figure S3.1). The amplicon was cloned into pJET1.2, according to the manufacturer's instructions, to create the vector pJETGluB4g1.2. Subsequently, the full gene was reamplified with primers STOSER 125 and 126 (containing a *Not*I restriction site) to remove the *Kpn*I restriction site upstream of the glutelin B4 promoter. The amplicon was cloned into pJET1.2, according to manufacturer's instructions, which resulted in the vector pJETGluB4g1.2 + *Not*I. This vector was digested with *Not*I and the restriction fragment containing the rice glutelin B4 gene was purified from an agarose gel. Simultaneously, the auxiliary vector pAUX3131 was digested with the same restriction enzyme and treated with shrimp alkaline phosphatase (SAP). The digested vector and the glutelin B4 gene restriction fragment were ligated to create pGluB4g31. In parallel, rice glutelin B4 cDNA was amplified from rice cDNA with primers STOSER 123 and 124 and cloned into pJET1.2 to create the vector pJETGluB4c1.2 (Supplemental figure S3.1). The vector was digested with restriction enzymes *Kpn*I and *Xho*I and the fragment was purified from an agarose gel. Vector pGluB4g31 was digested with the same restriction enzymes, dephosphorylated with SAP and ligated with the glutelin B4 cDNA fragment, which resulted in the vector pGluB4c31. In order to add *Sma*I sites into the glutelin B4 coding sequence (CDS) for further cloning (see below), primers STOSER 127 and 128 were used for amplification. Similarly, primers STOSER 129 and 130 were used to add *Sma*I sites at the 5' and 3' ends of the CDS, to enable removal of the CDS fragment in further cloning steps. Both amplicons were digested with *Sma*I and self-ligated to create the vectors pGluB4c31*Sma*ICDS and pGluB4c31*Sma*IdeltaCDS respectively (Supplemental figure S3.1). The former consists of the rice glutelin B4 promoter, the glutelin B4 cDNA sequence with a *Sma*I restriction site inside the CDS, the latter contains the same promoter followed by the 5' and 3' UTRs of glutelin B4, without the CDS.

Cloning of the *Arabidopsis thaliana* CDS encoding β- carbonic anhydrase 2 under control of the rice glutelin B4 promoter

In order to create the fusion sequence of sFBP with the CDS encoding β-carbonic anhydrase 2 (β-CA2 CDS) from *Arabidopsis thaliana* (AF428428), full β-CA2 cDNA was amplified with primers CA beta 2 1 forward and reversed (Supplemental figure S3.2). The PCR product was ligated with pJET1.2, which resulted in the pJET1.2/CA beta 2 vector. Subsequently, the β-carbonic anhydrase 2 CDS was amplified from this vector with primers STOSER 131 and 132, which contain *Eco*RV restriction sites for further cloning, and ligated with pJET1.2 to create vector pJET1.2/CA beta 2 *Eco*RV. The latter was digested with *Eco*RV and the β-CA2 fragment was purified from an agarose gel. Simultaneously, vector pGluB4c31*Sma*IdeltaCDS was submitted to a restriction digest with *Sma*I and treated with SAP. The linearized vector and the isolated β-CA2 CDS fragment were ligated by blunt end cloning, which resulted in the vector pGluB4c31CAbeta2 (Supplemental figure S3.2). This vector contains the rice glutelin B4 promoter, 5' and 3' glutelin B4 UTRs and the *Arabidopsis thaliana* β-CA2 CDS.

Cloning of sFBP and its fusion sequences with rice glutelin B4 and *Arabidopsis thaliana* β-CA2 under control of the rice glutelin B4 promoter

The DNA sequence of a soluble, synthetic folate binding protein (sFBP), based on bovine FBP (DN512948) (with codon optimization for expression in rice (with the use of the Codon Usage Database (http://www.kazusa.or.jp/codon/))(CCC and CCG were changed into CCT for proline; CGG and CGC into AGG for arginine; CTA and CTG into CTT for leucine; AGT and TCC into TCT for serine; GGG into GGC for glycine; ACG into ACT for threonine and GGG into GCA for alanine) and with removal of the membrane anchorage signal), was obtained from GenScript. A total of three transgenes were created with sFBP for expression in rice endosperm: sFBP, a fusion sequence of sFBP and β-CA2 (CAFBP) and a fusion sequence of sFBP with GluB4 (GluB4FBP). pGluB4c31CAbeta2 was digested with AleI, which cuts approximately halfway the β-CA2 CDS (in frame), SAP treated and ligated with the sFBP fragment (Supplemental figure S3.3). This resulted in vector pGluB4-CAFBP, in which the fusion sequence between sFBP and the β-CA2 CDS is placed under the control of the same rice glutelin B4 promoter. pGluB4c31*Sma*IdeltaCDS was cut with *Sma*I, SAP treated and ligated with sFBP, which resulted in the vector pGluB4-sFBP, in which sFBP is placed under the control of the rice glutelin B4 promoter (Supplemental figure S3.3). In addition, pGluB4-GluB4FBP was created by ligation of the sFBP fragment with the *Sma*I restricted and SAP treated pGluB4c31*Sma*ICDS vector (Supplemental figure S3.3). Here, the fusion sequence of sFBP and GluB4 cDNA is controlled by the same aforementioned promoter.

Cloning *Arabidopsis thaliana* cDNA encoding mitochondrial and cytosolic FPGS under the glutelin B1 promoter

In order to increase folate polyglutamylation, the mitochondrial (mt) and cytosolic (ct) folylpolyglutamate synthase (FPGS) isoforms of *Arabidopsis thaliana* were chosen for overexpression in rice seeds. The coding sequence of mtFPGS (accession number: AJ271786) was amplified from *Arabidopsis thaliana* cDNA with primers mtFPGS forward gateway and mtFPGS reversed gateway (Table 3.1), which contain partial attB sites for GATEWAY cloning. The attB sites were completed in a second

PCR reaction with primers attB1 and attB2. The fragment was purified from an agarose gel and subjected to a GATEWAY BP reaction with pDONR201, which resulted in vector pmtFPGS201 (Supplemental figure S3.4). Subsequently, a GATEWAY LR reaction was performed with this vector and destination vector pGluB169 (Storozhenko *et al.*, 2007), which resulted in the expression vector pGluB1/mtFPGS, in which the mtFPGS CDS is placed under the control of the glutelin B1 promoter. Similarly, ctFPGS (accession number: AJ292545) was amplified from *Arabidopsis thaliana* cDNA with primers ctFPGS forward gateway and ctFPGS reversed gateway (Table 3.1), which contain partial attB sites. Again, the attB sites were completed in a second amplification with primers attB1 and attB2. This fragment was extracted from an agarose gel and participated in a GATEWAY BP reaction with pDONR201 to create pctFPGS201. The latter was subjected to a GATEWAY LR reaction with pGluB169 to obtain the expression vector pGluB1/ctFPGS, in which the ctFPGS CDS is placed downstream of the rice glutelin B1 promoter.

Cloning mtFPGS, ctFPS and three sFBP transgenes into rice transformation vectors

A total of eleven rice transformation vectors were created, each of them carrying different transgene combinations for folate enhancement in rice. These include *Arabidopsis thaliana* cDNAs encoding GTP cyclohydrolase I (G), under the control of the rice endosperm specific globulin promoter, and aminodeoxychorismate synthase (A), controlled by the rice glutelin B1 promoter, in combination with mtFPGS or ctFPGS (both controlled by the glutelin B1 promoter) and/or one of the three FBP transgenes (all driven by the rice glutelin B4 promoter). The previously described pMOD35hGA rice transformation vector (with *hygromycin phosphotransferase II* as a selectable marker) (Storozhenko *et al.*, 2007) was used to create five triple gene constructs (pMOD35hGAmtF, pMOD35hGActF, pMOD35hGAsFBP, pMOD35hGACAFBP and pMOD35hGAGluB4FBP) and six quadruple gene constructs (pMOD35hGAmtFsFBP, pMOD35hGAmtFCAFBP, pMOD35hGAmtFGluB4FBP, pMOD35hGActFsFBP, pMOD35hGActFCAFBP and pMOD35hGActFGluB4FBP) (Figure 3.1). The mtFPGS and ctFPGS expression cassettes were cloned into pMOD35hGA by traditional cloning using the I-*Ppo*I homing endonuclease. Similarly, I-*Sce*I was used to clone the FBP expression cassettes into pMOD35hGA by traditional cloning.

Table 3.1: List of primers used for cloning and for expression analysis by real-time quantitative PCR.

Primer	Sequence	Primer specifications
STOSER 121	ATCGGCGTCTACAGCGCAGGCATCATA	Cloning of rice glutelin B4 gene
STOSER 122	TGGCGTGGCCAAAAGGTTCGGATCT	Cloning of rice glutelin B4 gene
STOSER 123	GATCAACCAGCCCAAGTTTCCAATAA	Cloning of rice glutelin B4 cDNA
STOSER 124	CAGAACCGCCACAAAGTTTCACATACT	Cloning of rice glutelin B4 cDNA
STOSER 125	GCGGCCGCCAATTATTTTGGACATTATGGAGAGAACA	Cloning of rice glutelin B4 gene, removal of *Kpn*I and addition of NotI restriction sites
STOSER 126	GCGGCCGCATCTAGGAATTATGTGTTGAAAGGACTT	Cloning of rice glutelin B4 gene, addition of NotI restriction site
STOSER 127	TTTAGTCCCGGGTGCAACGGTTTAGATGAGAACTTCT	Cloning of rice glutelin B4 CDS, addition of *Sma*I restriction site
STOSER 128	TTTAATCCCGGGGTTGTTACCCGCCAATAAGAACTCC	Cloning of rice glutelin B4 CDS, addition of *Sma*I restriction site
STOSER 129	AGTCCCGGGTGATGTACTAATGAAATAGTATAGG	Removal of rice glutelin B4 CDS, addition of *Sma*I restriction site
STOSER 130	AATCCCGGGAGCTATTTGAGGATGTTATTGGAAA	Removal of rice glutelin B4 CDS, addition of *Sma*I restriction site
STOSER 131	GATATCATGGGAAACGAATCATATGAAGACGCCAT	Cloning of *Arabidopsis thaliana* β- carbonic anhydrase 2 CDS, addition of EcoRV restriction site
STOSER 132	GATATCTCATATAGAATGAACGGGGGAAATT	Cloning of *Arabidopsis thaliana* β- carbonic anhydrase 2 CDS, addition of EcoRV restriction site
CA beta 2 1 forward	CCTGCTTCAGCCACTTCAAACTTGA	Cloning of *Arabidopsis thaliana* β- carbonic anhydrase 2 cDNA
CA beta 2 1 reversed	GGTAGCGATGGTGATGGTGATGTGT	Cloning of *Arabidopsis thaliana* β- carbonic anhydrase 2 cDNA
mtFPGS forward gateway	AAAAAGCAGGCTCCCTGATGCTCGTTTGTGGGAAAG	Cloning of *Arabidopsis thaliana* mtFPGS CDS, addition of partial attB sites
mtFPGS reversed gateway	AGAAAGCTGGGTTCATCTCTTTAGCAACCTGA	Cloning of *Arabidopsis thaliana* mtFPGS CDS, addition of partial attB sites
ctFPGS forward gateway	AAAAAGCAGGCTATCCTATGGCAACTGAAGACGATGGTGAA	Cloning of *Arabidopsis thaliana* ctFPGS CDS, addition of partial attB sites
ctFPGS reversed gateway	AGAAAGCTGGGTTCATTTCTTGATAAATCTCA	Cloning of *Arabidopsis thaliana* ctFPGS CDS, addition of partial attB sites
attB1	GGGGACAAGTTTGTACAAAAAAGCAGGCT	attB1 site for GATEWAY cloning
attB2	GGGGACCACTTTGTACAAGAAAGCTGGGT	attB2 site for GATEWAY cloning
STOSER 154	ATTATCACCTGAAGTTGAC	expression analysis GTPCHI
STOSER 155	GTGTAGCAATGAGTTCTT	expression analysis GTPCHI
STOSER 156	TGATTGTTGACCTTCTAA	expression analysis ADCS
STOSER 157	ACTGTTGTGTATGATTCT	expression analysis ADCS
STOSER 162	TACATCAAGAACCTCTCC	expression analysis rice tumor protein homologue (LOC_Os11g43900.1), reference gene
STOSER 163	ACCAACAAAGAACTGAAG	expression analysis rice tumor protein homologue (LOC_Os11g43900.1), reference gene
STOSER 164	GAACATGGAGAAGAACAA	expression analysis expressed protein (LOC_Os07g02340.1), reference gene
STOSER 165	CATATCTTGCACTGGATG	expression analysis expressed protein (LOC_Os07g02340.1), reference gene
STOSER 180	AATGAAATCTGGTCTCACTCT	expression analysis FBP
STOSER 181	CGAACCACATCTGAATGC	expression analysis FBP
mtFPGS 3 F	GGTCACTTCCAGTGCCGTAA	expression analysis mtFPGS
mtFPGS 3 R	AGCAACCTGAGCACATCTCC	expression analysis mtFPGS
ctFPGS 2 F	ACATTTGCGGAGTCTATTCTTCGTTG	expression analysis ctFPGS
ctFPGS 2 R	TCAACAGCCACAGGAAGTGACGAGAA	expression analysis ctFPGS

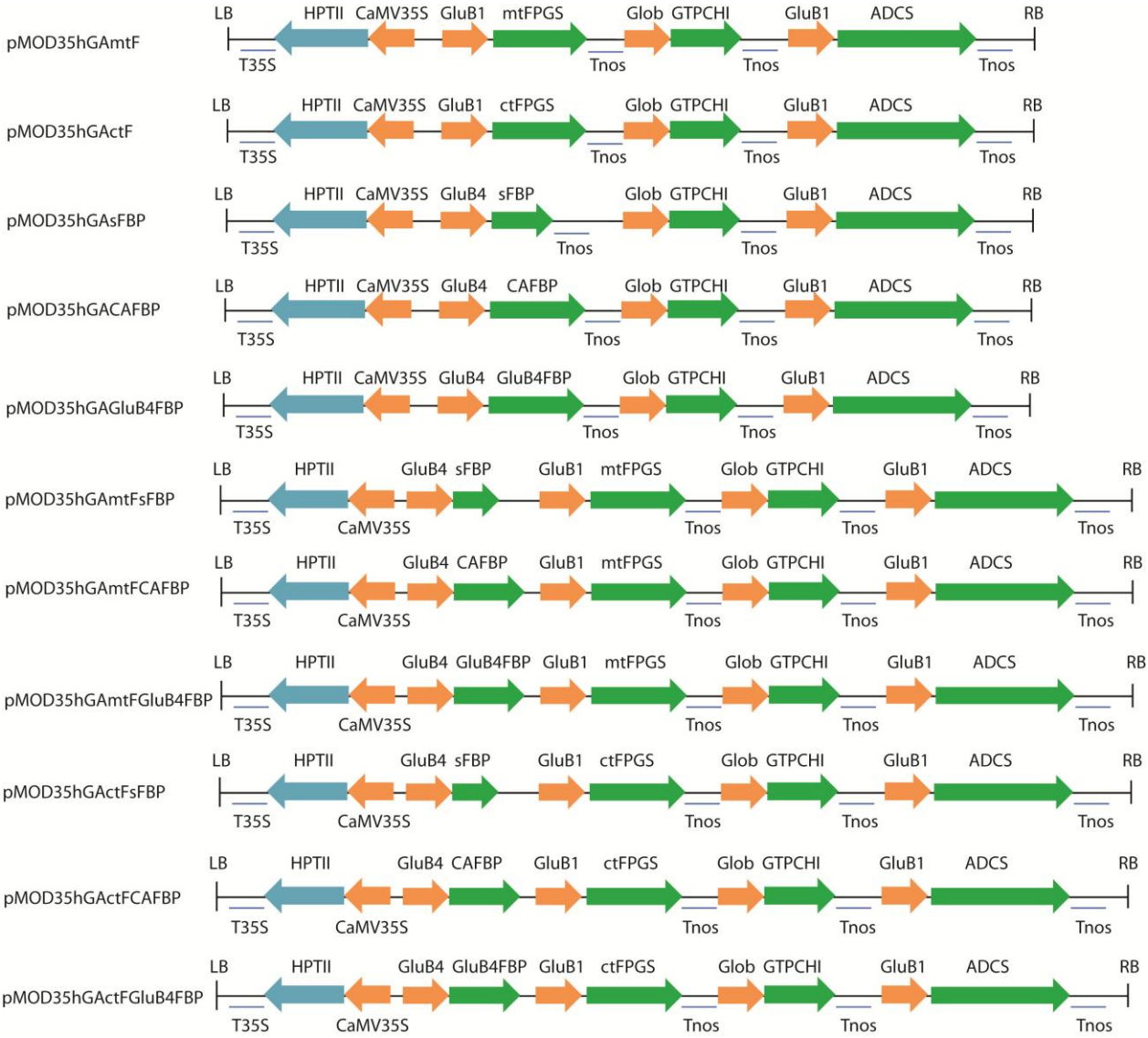


Figure 3.1: T-DNA vectors used for rice transformation. Orange arrows, promoters; green arrows, transgenes for folate enhancement and stability; blue arrows, hygromycin resistance gene (HPTII). Blue bars indicate transcriptional terminators. Abbreviations: LB and RB, left and right T-DNA borders; CaMV35S, core cauliflower mosaic virus 35S promoter; GluB1, rice glutelin B1 promoter; GluB4, rice glutelin B4 promoter; Glob, rice globulin promoter; GTPCHI, *Arabidopsis thaliana* cDNA encoding GTP cyclohydrolase I; ADCS, *Arabidopsis thaliana* cDNA encoding aminodeoxychorismate synthase; mtFPGS, *Arabidopsis thaliana* coding cDNA encoding mitochondrial folylpolyglutamate synthase (FPGS); ctFPGS, *Arabidopsis thaliana* cDNA encoding cytosolic FPGS; sFBP, soluble folate binding protein (FBP); CAFBP, fusion sequence between coding sequence of β- carbonic anhydrase 2 from *Arabidopsis thaliana* and sFBP; GluB4FBP, fusion sequence between rice glutelin B4 coding sequence and sFBP; HPTII, hygromycin phosphotransferase II; T35S, 35S transcriptional terminator; Tnos, nopaline synthase transcriptional terminator.

Microbial strains, plant material and transformations

Escherichia coli strain DH5α was used for all cloning steps, plasmid manipulations and propagation of all rice transformation vectors used for these experiments. *Agrobacterium tumefaciens* strain LBA4404 was used for T-DNA delivery from the binary vectors to rice. *Oryza sativa* subsp. *japonica* var. Nipponbare was soil-grown in short day light regime (8 hours of light (420 μmoles/m^2/s light intensity), 16 hours of darkness) at a constant humidity of 80% and 28°C and 21°C in light and darkness, respectively. Somatic embryogenic rice calli were propagated as previously described (Rueb *et al.*, 1994) and used for *Agrobacterium tumefaciens* mediated transformation (Scarpella *et al.*, 2000). All transformations were preformed twice to attempt to obtain a sufficient amount of transgenics, which are labelled as GAX(X) and 2GAX(X) (X being cytosolic or mitochondrial FPGS and/or sFBP (or its fusion sequences). Hygromycin was used to select transformed rice calli on solid media. Genomic DNA was isolated from rice leafs using the Invisorb Spin Plant Mini Kit (Invitek), according to the manufacturer's protocol, and used to select single copy T-DNA insertion events in the T0 generation, homozygous T-DNA status in the T1 generation and confirmation of this homozygosity in T2. Selection was done by real-time quantitative PCR using TaqMan probes, *hptII* specific primers and primers for rice sucrose phosphate synthase gene (as an internal reference), as previously described (Storozhenko *et al.*, 2007). Real-time quantitative PCR was performed using the iCycler-IQ5 system (Bio-Rad). Reaction conditions and copy number calculations were as previously described (Storozhenko *et al.*, 2007).

RNA extraction, cDNA synthesis and real-time quantitative PCR

To investigate the expression of the transgenes, total RNA was extracted from 100 mg of green rice seeds (in the milky stage) from each line using Trizol reagent (Invitrogen), according to the manufacturer's instructions, with minor modifications (1.5 ml of Trizol was used in each extraction and all reagents were scaled-up 1.5 times accordingly). The resulting samples were further purified with the RNeasy Plant Mini Kit (Qiagen), according to the manufacturer's instructions, with minor modifications (the RNA pellet obtained after Trizol purification was dissolved in buffer RLC and again, a volume up-scaling of 1.5 fold was performed for all reagents). All samples were treated with DNase (RNase-free DNase, Qiagen). Subsequently, 1 μg of rice seed RNA was converted into cDNA with the aid of the Verso cDNA Kit (Thermo-Scientific), using oligo-dT primers. The resulting cDNA was diluted 10 times with deionized water. Real-time quantitative PCR was performed using iCycler-IQ5 (Bio-Rad). KAPA Sybr® Fast qPCR Kit was used to investigate the expression of all transgenes. Reactions were carried out in a total volume of 25 μl, which consisted of 5μl of diluted cDNA, 12.5 μl of Sybr® Fast qPCR mix and 400nM (for GTPCHI, FBP, ctFPGS and expressed protein (LOC_Os07g02340.1)), 800nM (for ADCS and mtFPGS) or 200nM (for rice tumor protein homologue (LOC_Os11g43900.1)) of each primer. A real-time quantitative run comprised initial denaturation at 95°C for 3 minutes and 40 cycles of 15 seconds of denaturation at 95°C, 30 seconds of annealing at 55°C (for ADCS), 58°C (for FBP, mtFPGS and ctFPGS) or 60°C (for GTPCHI and both reference genes) and 30 seconds of extension at 72°C. Each reaction included two technical repeats. Data analysis and normalization was executed using the qBASE software, which is based on the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001; Hellemans *et al.*, 2007). Rice tumor protein homologue (LOC_Os11g43900.1) and expressed protein (LOC_Os0702340.1) were chosen as reference genes. All primers used for the expression analysis of the transgenes are listed in Table 3.1.

Folate stability experiments

Freshly harvested, ripe and dehusked rice seeds were used for both stability experiments. These seeds were stored in the dark at -80°C, 4°C, room temperature and 28°C for the pilot stability experiment and at -80°C and 28°C for the second stability experiment. Two to four samples (each consisting of 10 seeds) for folate analysis of each line were taken at zero time point and each following month up until 8 months and 4 months for the pilot and second stability experiment respectively.

Folate analysis

All materials, chemicals, stock and working solutions used and prepared for folate analysis were as previously described (De Brouwer *et al.*, 2010; Navarrete *et al.*, 2012). Sample preparation, UPLC™ conditions, mass spectrometric instrumentation and settings were as recently reported (Blancquaert *et al.*, 2013b).

Results

Transgenic rice seeds overexpressing GTPCHI, ADCS and mitochondrial or cytosolic FPGS

A first approach to improve folate stability in biofortified rice seeds consisted of the endosperm-specific overexpression of mitochondrial or cytosolic FPGS, both under the control of the rice glutelin B1 promoter. In two rounds of transformation, rice was transformed with the constructs pMOD35hGAmtF or pMOD35hGActF (Figure 3.1), both carrying the necessary transgenes for folate enhancement (*Arabidopsis thaliana* cDNA encoding GTPCHI (G) and ADCS (A) respectively). In the T0 generation, real-time quantitative PCR revealed 18 and 6 single-copy T-DNA transformants for GAmtF and GActF respectively. In the T1 and/or T2 generation (depending on whether homozygous lines were present in the T1 generation), 9 homozygous transgenic lines, originating from independent transformation events, could be detected for GAmtF and 2 for GActF (Supplemental table T3.1).

Folate analysis was used as a first screen for lines with enhanced folate content and polyglutamylation (Figure 3.2A and 3.2B) (Supplemental table T3.1). For GAmtF, 5 independently transformed lines (from 8 analysed) were found with folate levels above 100 µg/ 100 g fresh weight (FW). The highest folate content was found in line GAmtF 1.1.11 (1792.83 ± 772.17 µg / 100 g FW). Four lines were found with enhanced polyglutamylation (> 20%) in a folate enriched background: lines GAmtF 1.1.11, 19.1.3, 21.2.7 and 15.3.6. A maximum of 74% (74.11 ± 0.16 %) of polyglutamates was detected in line GAmtF 19.1.3. For GActF, no lines with significant folate enhancement (> 100 µg/ 100 g FW) could be obtained (from 2 independent lines analysed), the highest folate level being found in line GActF 3.2.4 (70.89 ± 3.05 µg/ 100 g FW). Therefore, these lines were not included in further experiments.

Transgenic rice seeds overexpressing GTPCHI, ADCS and FBP

A second strategy to improve folate stability in folate enhanced rice seeds, was through the overexpression of folate binding proteins (FBP), under the control of the rice glutelin B4 promoter. Three constructs were made (Figure 3.1): pMOD35hGAsFBP (overexpression of synthetic, soluble FBP), pMOD35hGACAFBP (overexpression of sFBP fused with *Arabidopsis thaliana* β-carbonic anhydrase 2)

and pMOD35hGAGluB4FBP (overexpression of sFBP fused with rice glutelin B4). All three constructs contained GTPCHI and ADCS for endosperm-specific folate enhancement. The approach to overexpress sFBP as a fusion sequence was chosen to enhance expression, a strategy which was successful in the accumulation of Japanese cedar pollen allergen Cryj1 in a fusion with glutelin B4 in rice seeds (Yang *et al.*, 2007). *Arabidopsis thaliana* β carbonic anhydrase 2 was chosen as a fusion partner because of the high abundance of this protein in the cytosol, rice glutelin B4 was chosen because it is one of the most abundant protein in rice endosperm. In the T0 generation, 27, 9 and 39 single-copy T-DNA transformed lines could be detected by real-time quantitative PCR for GAsFBP, GACAFBP and GAGluB4FBP respectively. In the T1 and/or T2 generation of GAsFBP, 9 homozygous transgenic lines, originating from independent transformation events were obtained, 3 for GACAFBP and 2 for GAGluB4FBP (Supplemental table T3.1).

Folate analysis was again used as a first screen for transgenic lines with enhanced folate content (Figure 3.2C) (Supplemental table T3.1). For GAsFBP, 5 independently transformed lines (out of 7 analysed) were found with folate levels higher than 100 µg /100 g FW. A maximum of 2066.85 ± 19.52 µg folate / 100 g FW was measured in line GAsFBP 2.2.5.7. An enhanced folate polyglutamylation was observed in two lines (GAsFBP 12.1.9.8 (50.08 ± 1.75% polyglutamylation) and GAsFBP 17.2.4 (68.90 ± 4.81 % polyglutamylation). For GACAFBP, all three lines examined (of which 2 originate from independent transformation events) showed an enhanced folate content. Line GACAFBP 3.1.3.2 contained the highest amount of folates (2529.78 ± 141.01 µg per 100 g FW). GACAFBP 3.1.3 showed the highest percentage of folate polyglutamates (31.42 ± 3.29 %) in combination with a high folate content (1172.29 ±263.70 µg/ 100 g FW). For GAGluB4FBP, 1 independent transgenic line (from 2 analysed) contained high folate levels (up to 1860.30 ± 206.62 µg/ 100 g FW in line GAGluB4FBP 17.2.4.6). However, folate polyglutamylation was not increased in those lines.

Transgenic rice seeds overexpressing GTPCHI, ADCS, mitochondrial or cytosolic FPGS and FBP

In a final approach to increase folate stability in metabolically engineered rice seeds, the two aforementioned strategies, i.e. increasing folate polyglutamylation and complexation with FBPs, were combined. This resulted in 6 rice transformation vectors, each of them carrying, besides GTPCHI and ADCS, one of the two FPGS isoforms and one of the three FBP constructs (Figure 3.1). In the T0 generation of transgenics, 34, 24 and 10 single-copy T-DNA transformants were detected for GAmtFsFBP, GAmtFCAFBP and GAmtFGluB4FBP respectively. In the T1 and/or T2 generation, 8 homozygous transgenic lines (resulting from independent transformation events) were obtained for GAmtFsFBP, 4 for GAmtFCAFBP and 5 for GAmtFGluB4FBP (Supplemental table T3.1).

Folate analysis showed two lines (from 5 independently transformed lines analysed) with an enhanced folate content for GAmtFsFBP (Figure 3.2D) (Supplemental table T3.1). Line GAmtFsFBP 14.1.5.3 had the highest folate content (596.81 ± 76.96 µg/ 100 g FW). An increased folate polyglutamylation (> 20% polyglutamates) was measured in line GAmtFsFBP 32.1.2 (42.38 ± 10.67 %). For GAmtFCAFBP, only 1 line (out of 4 independently transformed lines analysed) contained a folate level higher than 100 µg/ 100 g FW (GAmtFCAFBP 17.2.6.5 (106.16 ± 4.07 µg/ 100 g FW)). This line also showed an increased level of polyglutamylation (66.03 ± 3.67 %). For GAmtFGluB4FBP, again 1 line was obtained with enhanced folate

content (GAmtFGluB4FBP 44.1.14; 511.72 ± 83.67 µg/ 100 g FW), but folate polyglutamate abundance was not increased. For the constructs combining cytosolic FPGS with FBP, 2 single copy T-DNA insertion transgenics in the T0 generation were obtained for GActFsFBP, of which 1 homozygous line was detected in the T2 generation. For GActFCAFBP, 7 T0 transgenic lines with 1 copy of the T-DNA were found and 2 homozygous lines were obtained in the subsequent generation. Unfortunately, transformation with pMOD35hGActFGluB4FBP did not result in hygromycin resistant plants. For GActFsFBP, 3 independently transformed lines with enhanced folate content were obtained (from 3 analysed). The highest folate level was measured in line GActFsFBP 6.1.7 (1704.14 ± 22.70 µg/ 100 g FW), the highest folate polyglutamate abundance in line 4.2.23 (37.70 ± 1.97 % polyglutamates). For GActFCAFBP, 2 independently transformed folate biofortified lines were observed (from 3 analysed). Line 2GActFCAFBP contained 1673.07 µg folate per 100 g FW. No increase in the ratio poly/monoglutamatylated folate was found in those lines.

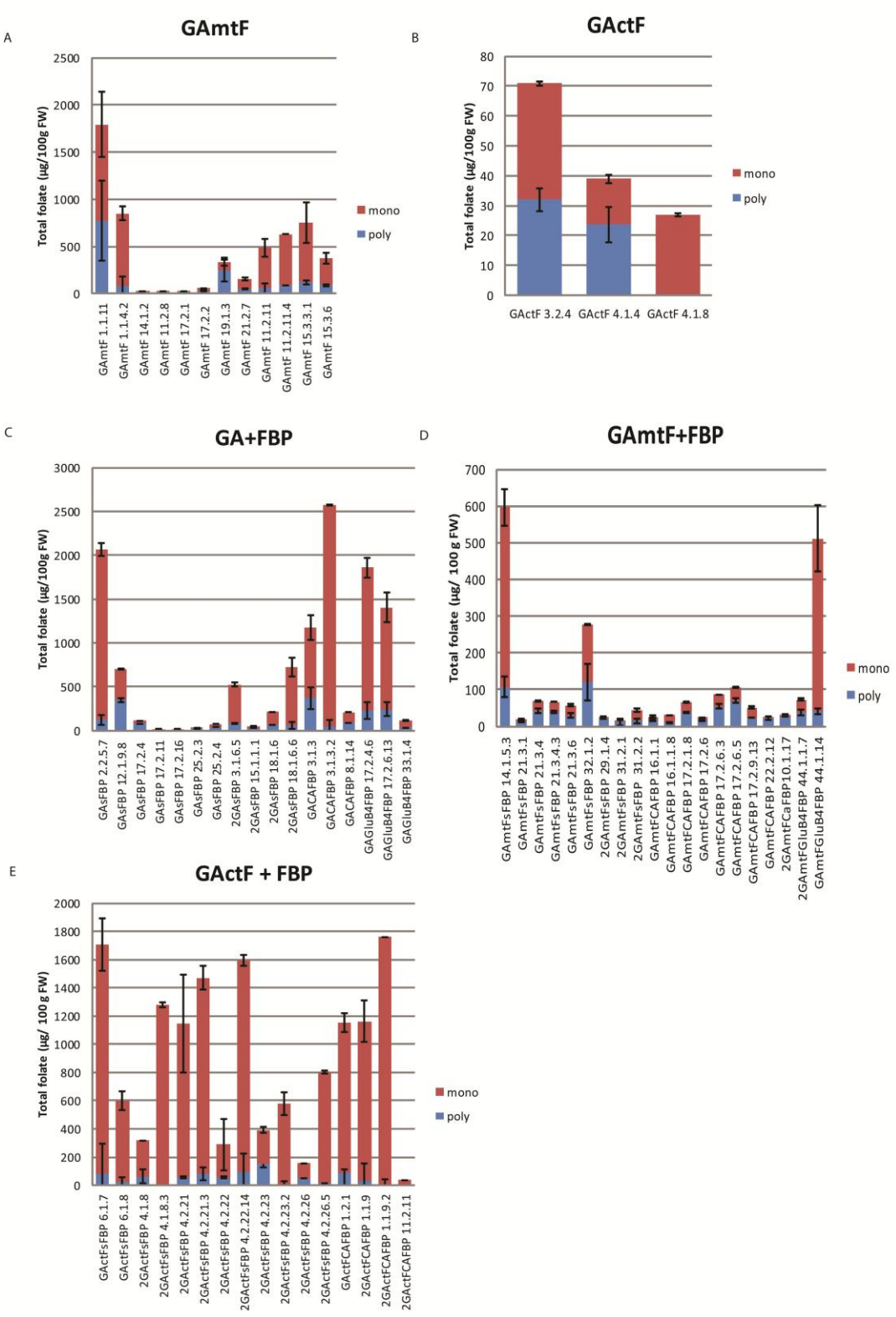


Figure 3.2: Total folate content of transgenic rice seeds, transformed with the vectors shown in figure 1. Values are means of two independent samples; error bars indicate standard deviation. The folate monoglutamate fraction is represented by red bars, the polyglutamate fraction by blue bars.

Pilot stability experiment with wild type and folate enhanced rice seeds

In order to investigate folate stability in wild type and folate biofortified rice seeds, WT (43.21 ± 1.22 µg folates / 100g FW; 73% polyglutamylation) and two GA lines (GA9.15 (716.47 ± 147.45 µg folates / 100g FW; 13% polyglutamylation) and GA26.5 (559.71 ± 33.63 µg folates / 100g FW; 16% polyglutamylation)) were selected for participation in a pilot stability experiment. Freshly harvested, mature seeds were dehusked and stored at various temperature conditions (-80 °C, 4°C, room temperature (RT) and 28°C). Samples were analysed on folate content on a monthly basis, up until 8 months (with the exception of month 7), (Figure 3.3). In all three lines, after 8 months of storage, approximately 60% of folate was degraded at 4°C, RT and 28°C. At -80°C, folate levels in GA9.15 remained unaltered after 8 months, while decreasing 50% in WT and approximately 40% in GA26.5. At non-freezing conditions, folate levels in all three lines raised after 1 month of storage, except at RT, where folate content dropped continuously. Degradation continued until month 4 at 4°C and 28°C, from when folate levels stabilized upon further storage. At RT, folate content remained unaltered after 3 months of storage.

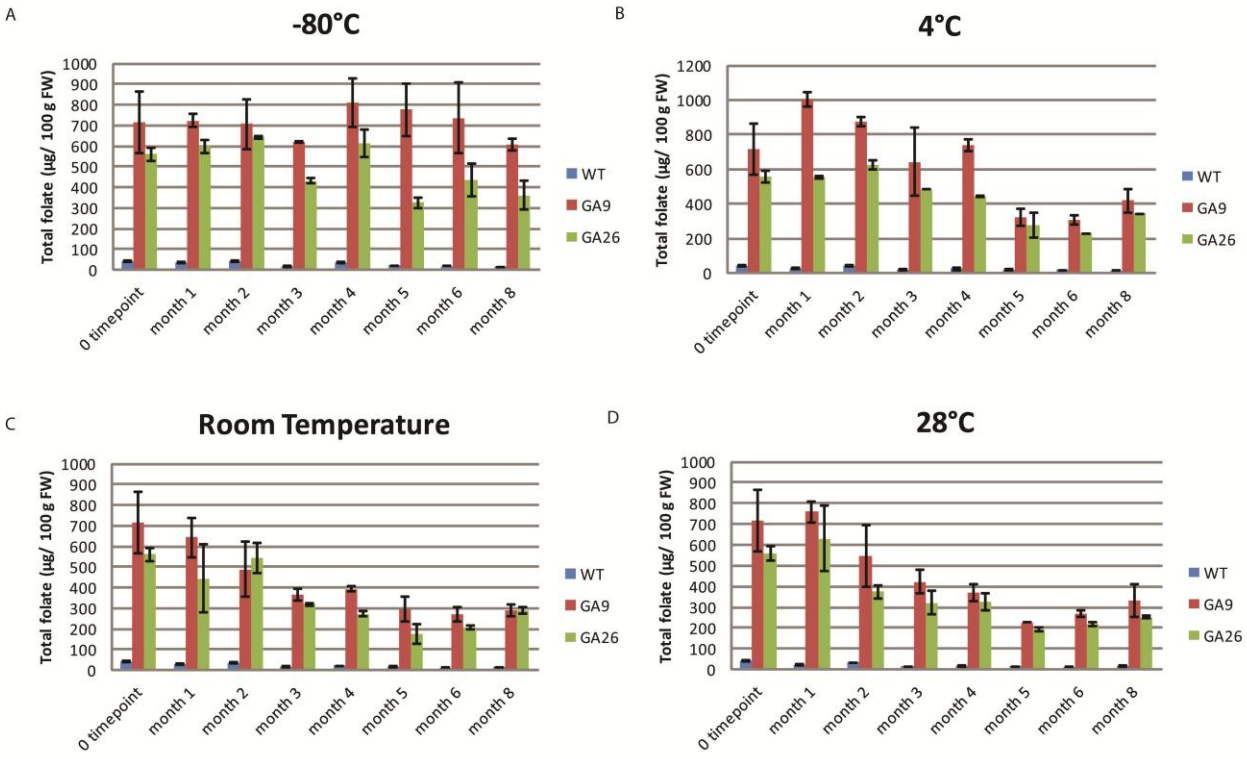


Figure 3.3: Total folate levels in WT, GA9.15 and GA26.5 up until 8 months of storage at various temperature conditions. Values are means of two independent samples; error bars indicate standard deviation.

Stability experiment with folate enhanced and stability engineered rice seeds

In order to investigate whether folate polyglutamylation and/or folate complexation with folate binding proteins would result in enhanced folate stability, a second, more elaborate folate stability experiment was conducted. A total of 27 lines were included in this experiment: WT, GA9.15 and GA26.5 were

included as control lines, together with 24 homozygous lines (from T2 and T3 generation) containing different gene combinations. These comprised 5 GAmtF (from 4 independent transformation events), 3 GAsFBP (3), 3 GACAFBP (3), 2 GAGluB4FBP (1), 2 GAmtFsFBP (2), 2 GAmtFCAFBP (2), 2 GAmtFGluB4FBP (2), 4 GActFsFBP (2) and 1 GActFCAFBP line. Analysis of parental lines showed an enhanced folate content in all cases (Figure 3.2). Since no difference in folate degradation rate could be observed after 8 months of storage at temperatures above 0°C during the pilot stability experiment (Figure 3.3), 28°C was chosen as the main temperature condition to study folate stability in the new engineered lines, a condition at which seeds are stored in rice consuming countries. Additionally, since folate degradation stabilized after 4 months at 28°C (Figure 3.3D), it was decided to restrict this stability experiment to 4 months. Although the results of the pilot folate stability experiment remained inconclusive about folate degradation at -80°C (Figure 3.3A), this temperature condition was included as a reference point. Unfortunately, due to loss of all samples of the time points zero and first month during folate analysis, folate levels of all samples, stored at -80°C, of months 2, 3 and 4 were averaged and considered as zero time point for each individual line. The results are shown in figure 3.4. In general, folate levels in all lines examined was lower than that in their parental lines. For GAmtF, 3 lines show a decrease in folate content between 25 and 45% as compared to the zero time point (Figure 3.4B) (percentages of folate degradation of all lines are listed in Supplemental table T3.2). One line (GAmtF 15.3.3.1) even showed an increase in folate content after 2 months of storage at 28°C. For GAsFBP, all three transgenic lines have a folate degradation between 21 and 37% after 4 months of storage. Folate content in GACAFBP lines have a drop in folate content of maximally 37%. However, no folate degradation could be observed in line GACAFBP 3.1.3.2, indicating an enhanced folate stability. In transgenics overexpressing FBP in fusion with glutelin B4 (GAGluB4FBP), a drop in folate level of maximally 12 % could be measured. For GAmtFsFBP, folate content dropped between 10 and 22%, the highest folate stability being found in line 2GAmtFsFBP 14.1.5.3. For GAmtFCAFBP, only 1 line (2GAmtFCAFBP 7.2.1) had an enhanced folate content, although folate levels at month 4 were 35 % degraded as compared to zero time point. The other line, 2GAmtFCAFBP 10.1.16, had folate levels within WT range. In both GAmtFGluB4FBP transgenic rice lines, folate levels remained stable after 4 months of storage at 28°C. For GActFsFBP, folate content maximally dropped 33% after 4 months and the highest folate stability was measured in line 2GActFsFBP 4.2.21.3, where folate levels remained stable. A stable folate content was also found in line GActFCAFBP 1.1.9.2. Unfortunately, the results previously obtained for WT and the two GA lines (Figure 3.3) could not be repeated in this experiment and even an increase in folate content was observed for line GA26.5. Folate levels in WT increased approximately 10 % after 4 months of storage, whereas in GA9.15 folate content dropped 8% (Supplemental table T3.2).

Folate polyglutamylation was investigated in all participating lines (Figures 3.5, 3.6, 3.7 and 3.8). The highest degree of polyglutamylation was found in WT, although the ratio poly/mono in this control line was lower than previously observed (ratio poly/mono was approximately 50 % in WT rice seeds, described by Storozhenko *et al.*, 2007). The occurrence of enhanced folate polyglutamates, observed in the parental generation of some of the transgenics (figure 3.2), could not be detected.

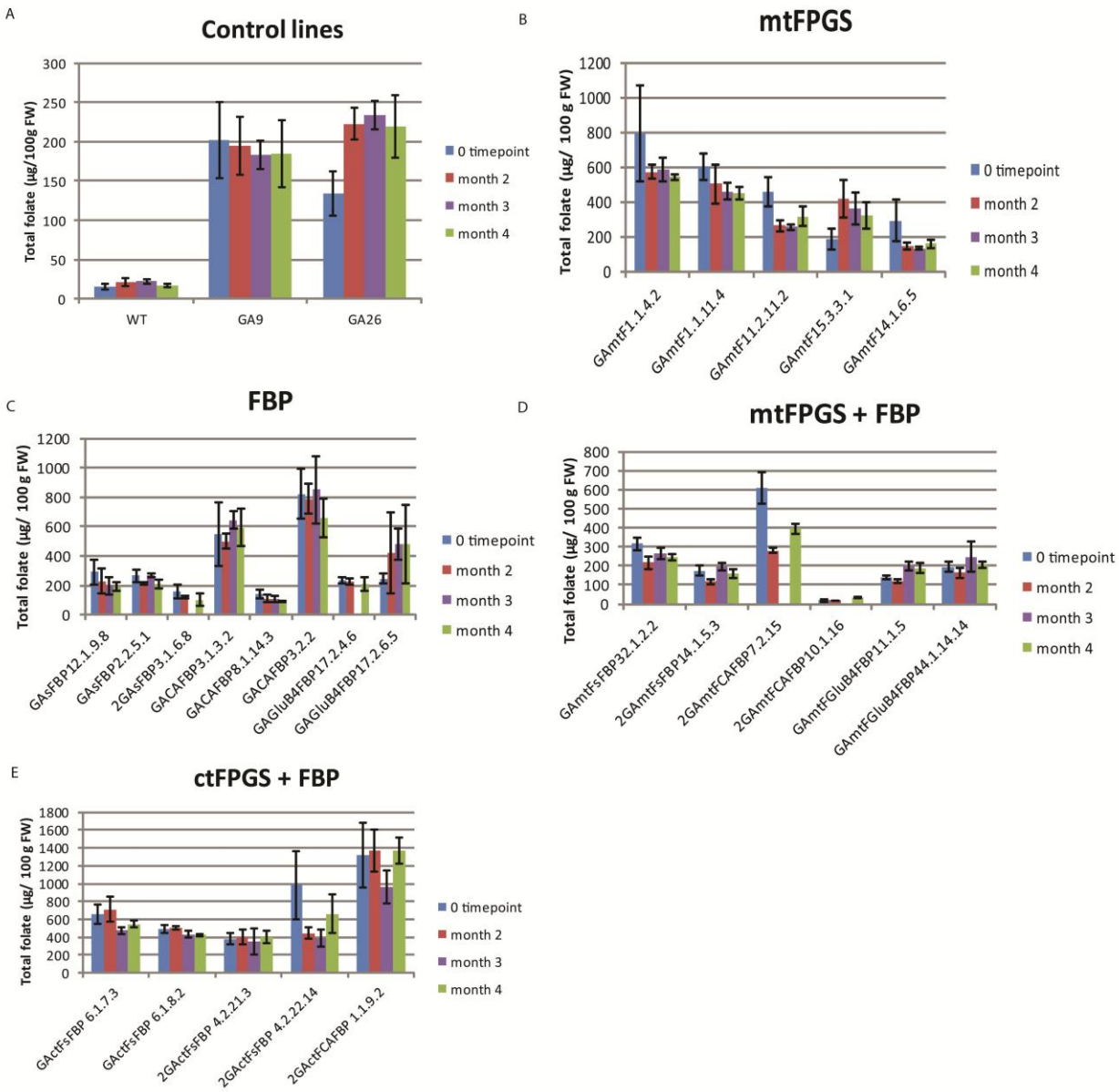


Figure 3.4: Total folate levels in rice seeds engineered for folate content and stability upon storage at 28°C. Zero time point was measured by calculating average folate content in samples of month 2, 3 and 4 of each line, kept at -80°C. Values are means of two independent samples for month 3 and four independent samples for months 2 and 4; error bars indicate standard deviation.

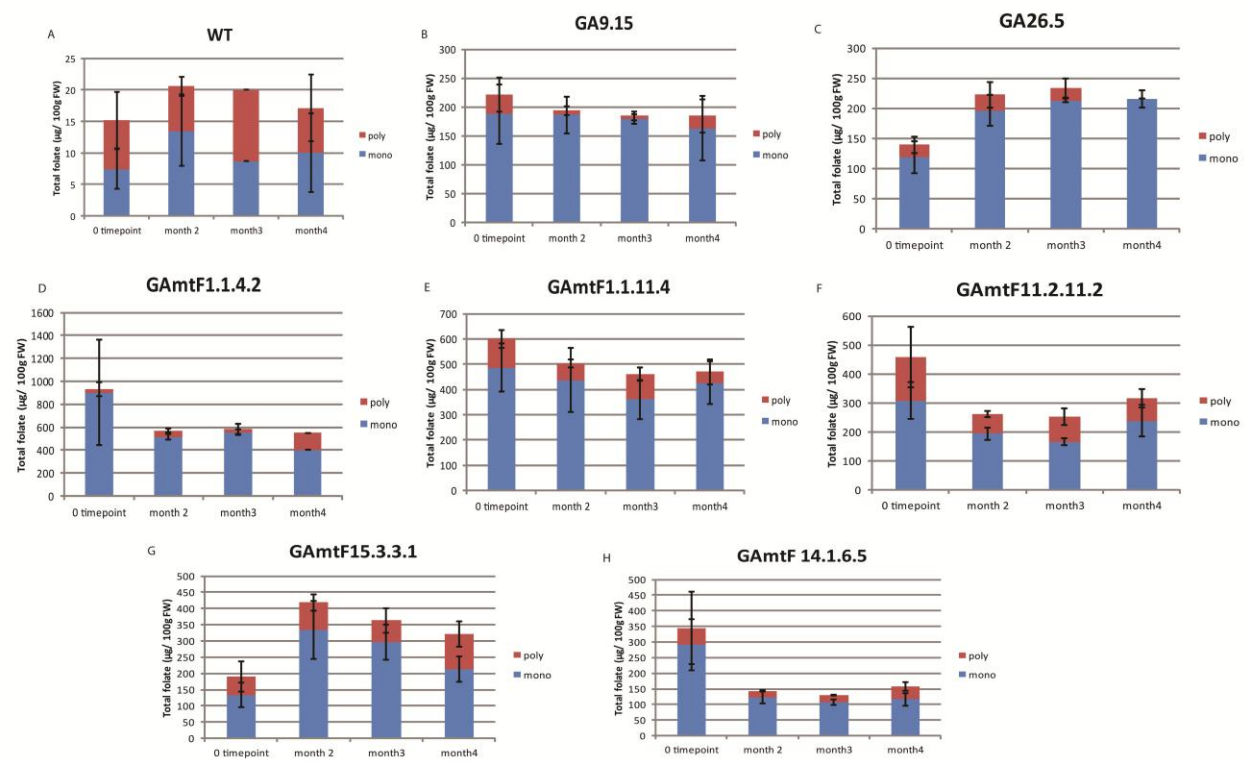


Figure 3.5: Effect of combined G, A, and mtFPGS overexpression on folate content and polyglutamylation. Distribution of folate mono- and polyglutamates in control lines (A,B,C) and rice seeds engineered in folate content and polyglutamylation (mtFPGS) (D,E,F,G,H) upon storage at 28°C. Zero time point was measured by calculating average folate content in samples of month 2, 3 and 4 of each line, kept at -80°C. Values are means of two independent samples for month 3 and four independent samples for months 2 and 4; error bars indicate standard deviation.

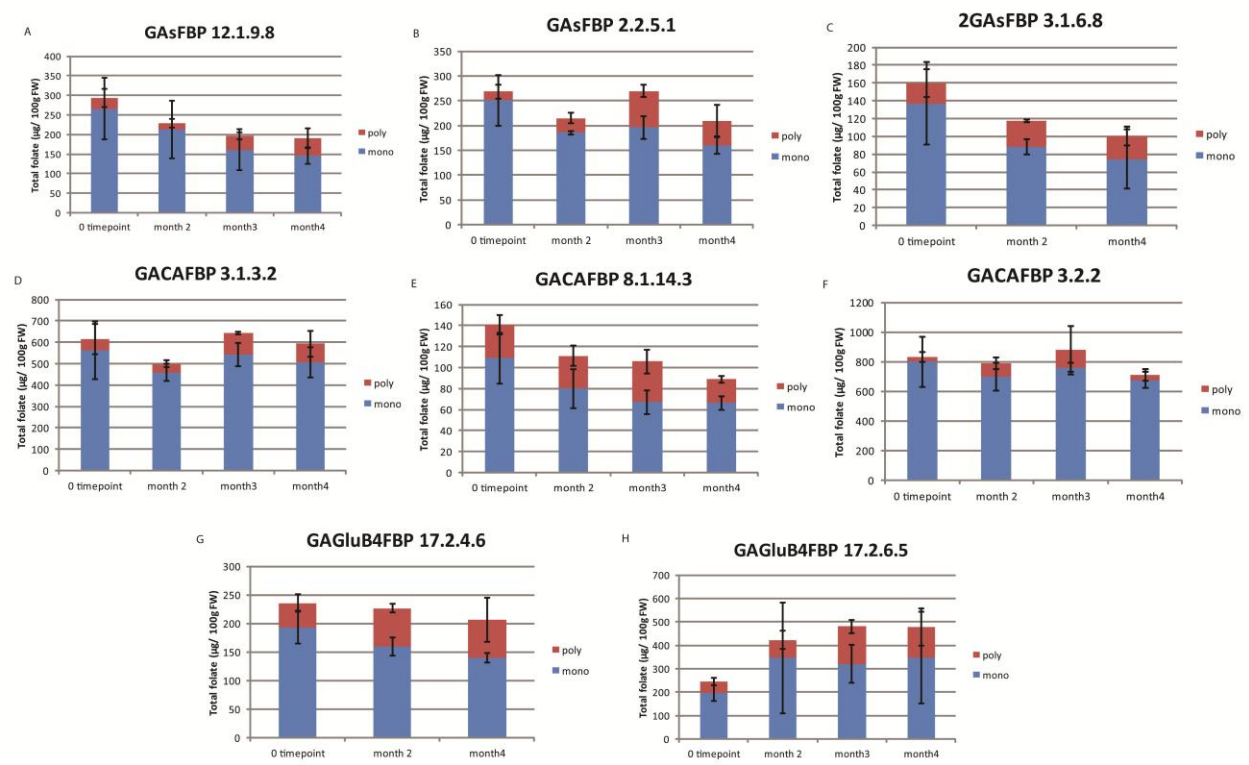


Figure 3.6: Effect of combined G, A, and FBP overexpression on folate content and polyglutamylation. Distribution of folate mono- and polyglutamates in rice seeds engineered in folate content and enhanced stability through complexation with folate binding proteins upon storage at 28°C. Zero time point was measured by calculating average folate content in samples of month 2, 3 and 4 of each line, kept at -80°C. Values are means of two independent samples for month 3 and four independent samples for months 2 and 4; error bars indicate standard deviation.

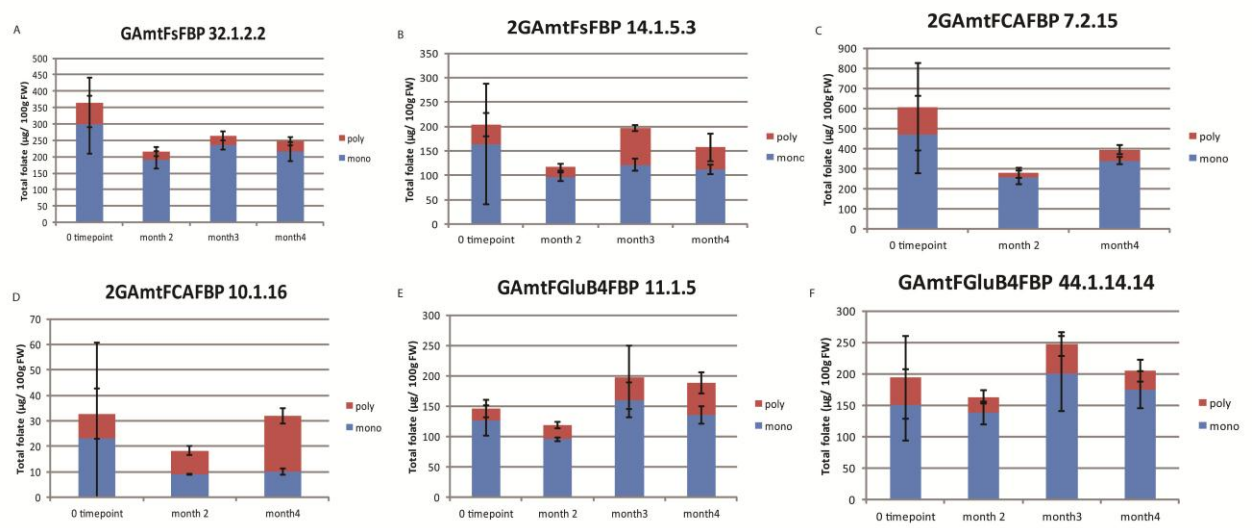


Figure 3.7: Effect of combined G, A, mtFPGS and FBP overexpression on folate content and polyglutamylation. Distribution of folate mono- and polyglutamates in rice seeds engineered in folate content and enhanced stability through polyglutamylation (mtFPGS) and complexation with folate binding proteins upon storage at 28°C. Zero time point was measured by calculating average folate content in samples of month 2, 3 and 4 of each line, kept at -80°C. Values are means of two independent samples for month 3 and four independent samples for months 2 and 4; error bars indicate standard deviation.

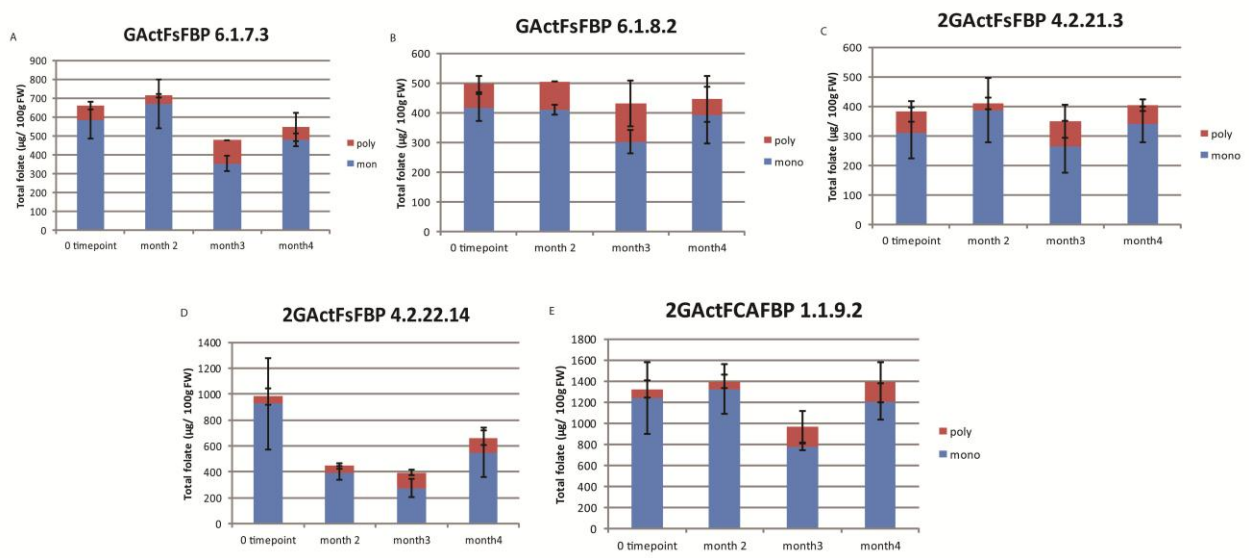


Figure 3.8: Effect of combined G, A, ctFPGS, and FBP overexpression on folate content and polyglutamylation. Distribution of folate mono- and polyglutamates in rice seeds engineered in folate content and enhanced stability through polyglutamylation (ctFPGS) and complexation with folate binding proteins upon storage at 28°C. Zero time point was measured by calculating average folate content in samples of month 2, 3 and 4 of each line, kept at -80°C. Values are means of two independent samples for month 3 and four independent samples for months 2 and 4; error bars indicate standard deviation.

Transgene expression in folate enhanced and stability engineered rice seeds

The expression of the transgenes in all lines which participated in the second stability experiment was investigated in the milky stage of rice seed development by real-time quantitative PCR (Figure 3.9). The level of expression of all transgenes varied greatly among the different lines examined (Supplemental table T3.3). GTPCHI overexpression could be detected in all transgenics, except for lines 2GAmtFCAFBP 7.2.15, GAmtFGluB4FBP 11.1.5 and GACAFBP 3.2.2 (Figure 3.9A). ADCS overexpression was absent in the same three transgenic lines and in line GAGluB4FBP 17.2.6.5 (Figure 3.9B). Cytosolic FPGS was expressed in all lines examined (Figure 3.9C), as well as mitochondrial FPGS (with the exception of line 2GAmtFCAFBP 7.2.15, where no overexpression could be detected) (Figure 3.9D). FBP overexpression was not observed in lines 2GAmtFsFBP 14.1.5.3, 2GAmtFCAFBP 7.2.15, GAmtFGluB4FBP 11.1.5, GACAFBP 3.1.3.2, GACAFBP 8.1.14.3 and GAGluB4FBP 17.2.6.5 (Figure 3.9E). Although GTPCHI and ADCS overexpression could not be measured in some transgenic lines, it is possible that it occurred earlier on in rice seed development, explaining why folate levels in the respective lines are still enhanced.

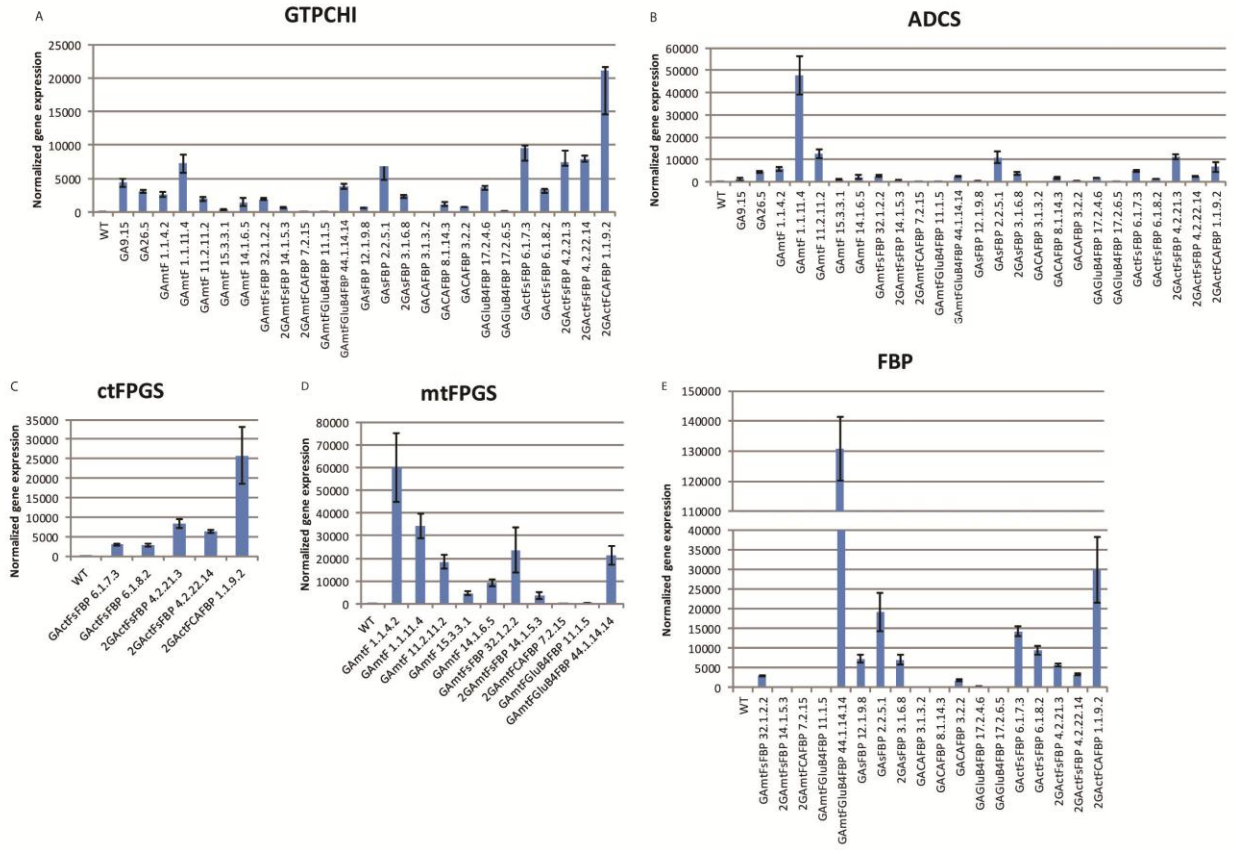


Figure 3.9: Expression levels of GTPCHI, ADCS, FPGS and FBP transgenes in green rice seeds of all lines participating in the second stability experiment. Expression analyses were performed by real-time quantitative PCR. Rice tumor protein homologue (LOC_Os11g43900.1) and expressed protein (LOC_OS07g02340.1) were used as reference genes for normalization. Values are means of three technical replicates; error bars indicate standard error. Data analysis was performed using qBASE, based on the $2^{-\Delta\Delta Ct}$ method.

Discussion

In order to enhance folate stability in biofortified rice, 11 transformation vectors were created (Figure 3.1). Each transformation vector contains the necessary genes for rice folate enhancement (Storozhenko *et al.*, 2007): GTPCHI, which catalyzes the first step in the pterin branch and ADCS, which catalyzes the first step in the p-ABA branch of the folate biosynthesis pathway. On each vector, these two cDNAs, originating from *Arabidopsis thaliana*, were combined with mitochondrial or cytosolic FPGS and/or one of the FBP forms. FPGS is responsible for the last step in the biosynthesis of tetrahydrofolate (THF) polyglutamate and adds glutamate units to the monoglutamylated form of THF. As a result, the anionic nature of the folate molecule is increased, which favors cellular retention and hence possibly folate stability. Complexation of folates with FBPs is a naturally occurring process to enhance folate stability. FBPs can be found in e.g. mother milk, stabilizing folates prior to consumption by the progeny.

By increasing polyglutamylation (GAmtF and GActF constructs), a higher abundance of polyglutamylated folates in transgenic rice transformed with GAmtF and GActF was obtained, as compared to the previously described GA lines (Storozhenko *et al.*, 2007). However, high levels of folate could not be obtained for the GActF lines, where folate levels were maximally augmented 3-fold as compared to the WT (Figure 3.2B). Since approximately 50 % of the folate pool occurs as polyglutamates in WT rice seeds (Storozhenko *et al.*, 2007), it is possible that the increase in abundance of polyglutamylated folates in these transgenics is caused by the activity of the endogenous rice FPGS. It is worth mentioning that folate enhancement by metabolic engineering in rice did not affect the expression of the endogenous folate biosynthesis genes (Blancquaert *et al.*, 2013b). This would mean that endogenous FPGS expression is sufficient to maintain a high poly/mono ratio when folate levels remain considerably low, as is the case for GActF transgenics. However, more transgenic lines of this construct need to be analyzed to put weight to this assumption. On the other hand, a couple of GAmtF transgenics could be obtained with folate levels up to those in earlier work (Storozhenko *et al.*, 2007) (Figure 3.2A). In line GAmtF 1.1.11 and other lines of this construct, an increase in abundance of polyglutamylated folates was observed. Since the mitochondria are the main site of THF biosynthesis, it is plausible that folates accumulate in this subcellular compartment in biofortified rice seeds. However, subcellular fractionation studies are necessary to prove this statement.

In a second approach to enhance folate stability in engineered rice seeds, several forms of FBP were overexpressed. Since high folate levels could be obtained by overexpressing GTPCHI and ADCS, a massive accumulation of FBP is required to bind the majority of those biosynthesized folates. Therefore, it was decided to overexpress FBP as such and as a fusion with rice glutelin B4 or *Arabidopsis thaliana* β-carbonic anhydrase 2 (Figure 3.1). These proteins are highly abundant in rice endosperm and the *Arabidopsis* cytosol, respectively. By creating fusion sequences, it was attempted to stabilize their expression in the endosperm, since plants have the ability to recognize and silence non-plant genes and the existence of FBPs in plants remains unknown. Folate analysis revealed that several transgenic lines contained folate levels much higher than those previously described (Storozhenko *et al.*, 2007) (Figure 3.2C), which indicates that FBP indeed binds to folates and sequesters them. The highest folate content

in all transgenics described in this study was found in line GACAFBP 3.1.3.2. However, high folate levels were measured in GAsFBP and GAGluB4FBP as well, indicating that fusion sequences are not necessarily needed to obtain a high abundance of FBP in rice endosperm.

Finally, the two aforementioned strategies were combined, which resulted in a total of 6 rice transformation vectors (Figure 3.1). The highest folate levels in those transgenics were found in combination with cytosolic FPGS (Figure 3.2D and E). However, no high increase in folate polyglutamates could be measured in those transgenics, which makes it unclear whether the combination of these two approaches would result in a higher folate stability as compared to each strategy by itself. Unfortunately, the final stability experiment remained inconclusive, due to loss of zero time point samples and samples of month 1. Repeating this experiment is required to gain full insight whether the aim of this study is reached or not.

Although high folate levels could be obtained in several transgenic lines described in this study, it is striking that in the progeny of those first generation transgenics, folate levels dropped. Since folates are important in gene methylation (Ly *et al.*, 2012), it is possible that high levels of folates promote gene silencing, especially in transcriptionally active loci, such as a T-DNA insertion, on which transgenes are controlled by strong promoters. The stacking of genes on a single T-DNA locus would only increase this effect. Therefore, retransformation with single or double gene constructs could offer a solution, although this would be more elaborate and time-consuming. Moreover, the latter would require the use of different selectable markers. Unfortunately, due to the difficulties in GMO legislation and regulation, this strategy could possibly hamper the acceptance of multigene engineered transgenics.

The pilot stability experiment clearly shows that folates are susceptible to degradation by long storage (Figure 3.3). However, since no real difference was observed in folate stability after 6 and 8 months of storage for the non-freezing temperature conditions, it can be concluded that temperature, at least at low and moderate temperatures, does not affect folate stability. Rice seeds stored at -80°C are most likely still prone to degradation, as confirmed in WT and GA26.5 (Figure 3.3A). Therefore, using average folate levels of samples kept at -80°C as zero time point reference, as is done for the second stability experiment, is incorrect. In some lines (e.g. GA9.15 at 4°C; Figure 3.3B), an increase in folate content could be measured after 1 month of storage. This can be explained, or at least partially, by the fact that dehydration still occurs in the first month of storage. This way, folate concentration increases as water evaporates out of the seeds. Indeed, seed weight dropped after zero time point (data not shown), clearly indicating that dehydration is occurring at that moment.

Conclusions

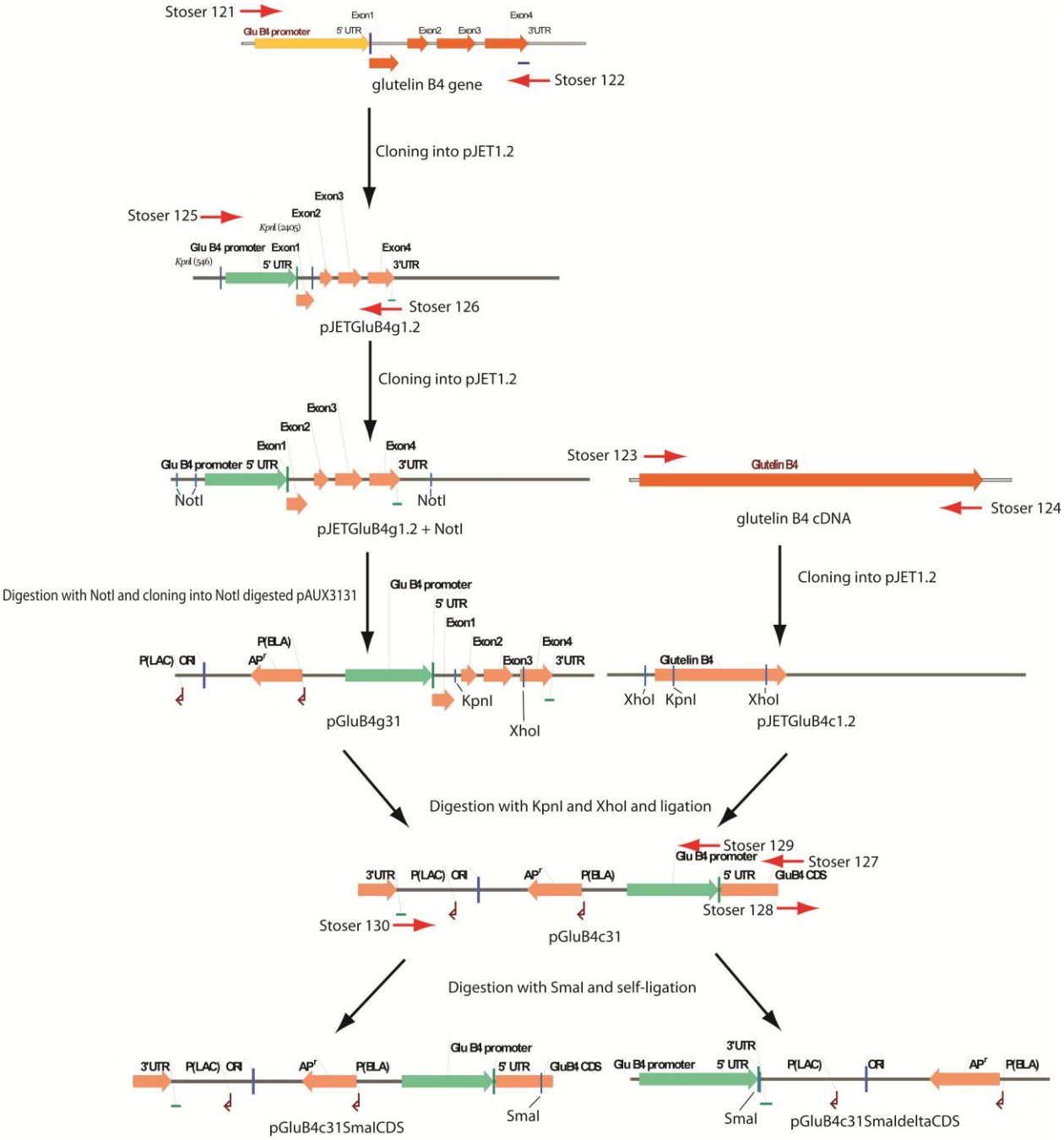
We conclude that overexpressing FBP in rice seeds, in combination with GTPCHI and ADCS, further increases folate content as compared to GA lines. This is most likely because complexation with FBP enhances sequestration. Since a stability experiment with all newly created transgenic lines needs to be repeated, unfortunately, no conclusions can be drawn about the ability of the two described strategies

to enhance folate stability in biofortified rice seeds at this point. Nevertheless, considering the fact that transgenic rice lines could be obtained with higher folate levels than reported for GA lines (Storozhenko *et al*., 2007) and that certain transgene combinations (such as GActFCAFBP and GAmtF) resulted in an apparently higher occurrence of folate enhanced transgenics, there are promising indications of an enhanced folate stability in these engineered rice seeds.

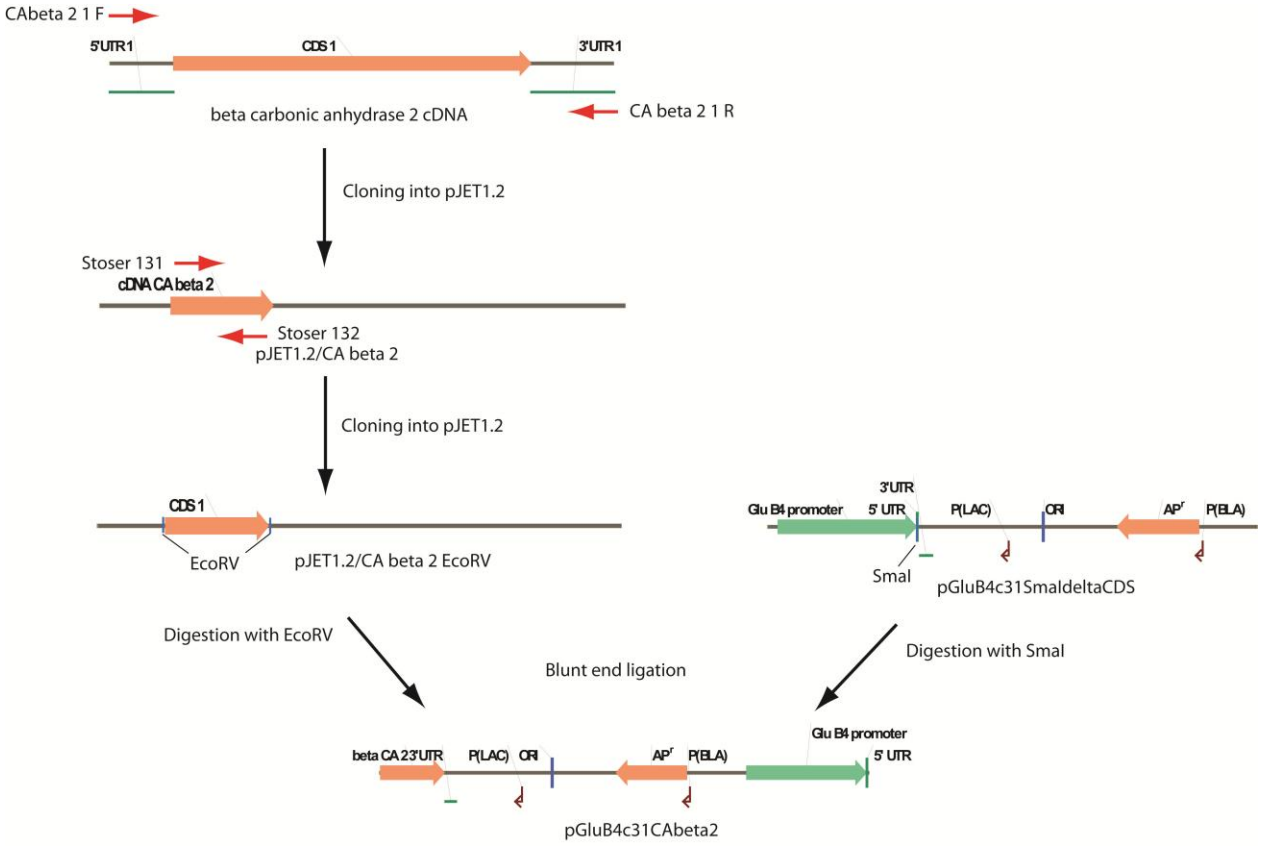
Acknowledgements

DVDS and WL gratefully acknowledge financial support from Ghent University (Bijzonder Onderzoeksfonds, BOF2004 /GOA/ 012 and BOF2009 /G0A /004), and the Fund for Scientific Research Flanders (FWO, project 3G012609).
D.B. is indebted to FWO (Fonds voor Wetenschappelijk Onderzoek – Vlaanderen), for a PhD grant.

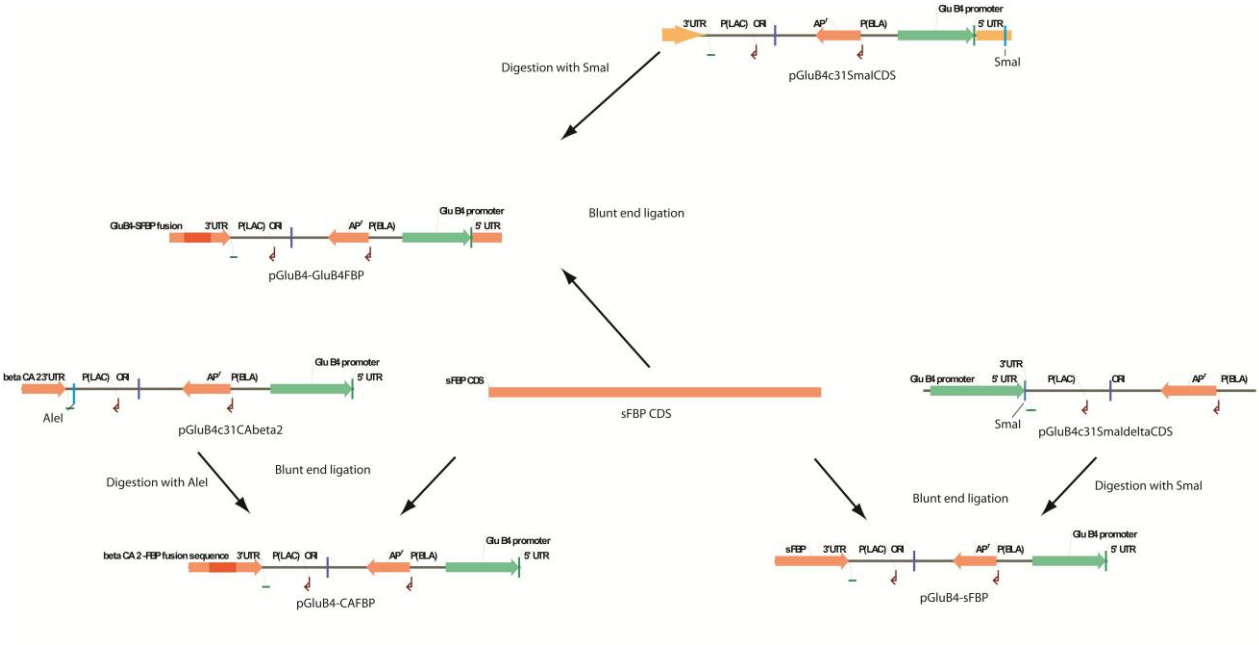
Supplementary figures and tables



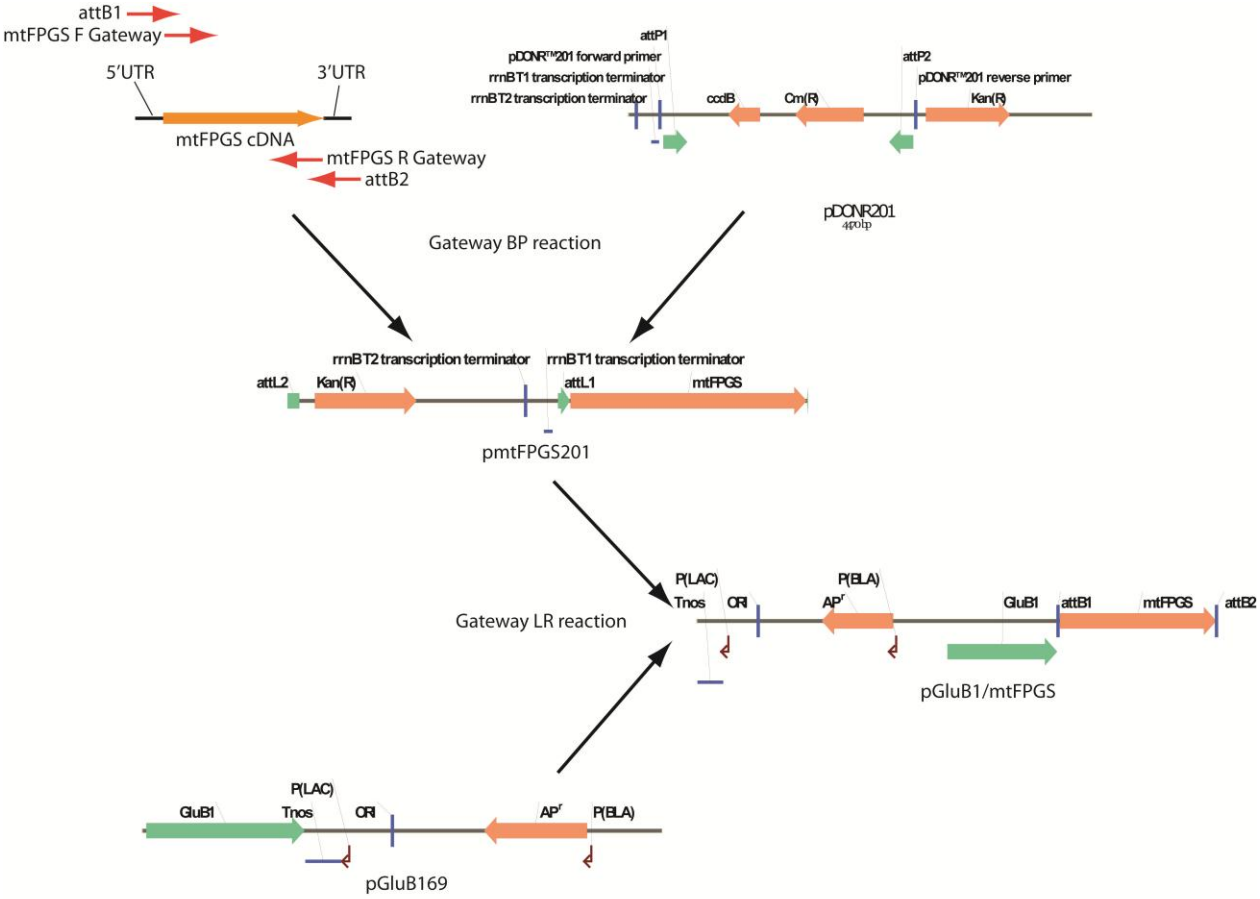
Supplemental figure S3.1: Schematic overview of the different steps followed for the cloning of the rice glutelin B4 promoter and cDNA and auxiliary vectors pGluB4c31SmaICDS and pGluB4c31SmaIdeltaCDS in which FBP was cloned in further reactions.



Supplemental figure S3.2: Schematic overview of the different steps followed for the cloning of *Arabidopsis thaliana* cDNA encoding β-carbonic anhydrase 2 into auxiliary vector pGluB4c31*Sma*IdeltaCDS to create vector pGluB4c31CAbeta2.



Supplemental figure S3.3: Schematic overview of the different steps followed for the cloning of sFBP into vectors pGluB4c31*Sma*IdeltaCDS, pGluB4c31*Sma*ICDS and pGluB4c31CAbeta2 to create vectors pGluB4-sFBP, pGluB4-GluB4FBP and pGluB4-CAFBP.



Supplemental figure S3.4: Schematic overview of the different steps followed for the cloning of mtFPGS by Gateway cloning. A similar strategy was followed for ctFPGS.

Constructs	Homozygous T1	Homozygous T2	number of high folate lines	Maximal % polyglutamylation
GAmtF	9	0	5 (8 analyzed)	74%
GActF	1	1	0 (2 analyzed)	/
GAsFBP	8	1	5 (7 analyzed)	69%
GACAFBP	2	1	2 (2 analyzed)	31%
GAGluB4FBP	2	0	1 (2 analyzed)	/
GAmtFsFBP	7	1	2 (5 analyzed)	42%
GAmtFCAFBP	4	0	1 (4 analyzed)	66%
GAmtFGluB4FBP	5	0	1 (1 analyzed)	/
GActFsFBP	1	0	3 (3 analyzed)	38%
GActFCAFBP	2	0	2 (3 analyzed)	/
GActFGluB4FBP	0	0	/	/

Supplemental table T3.1: Summary of all constructs used for this study, with the number of homozygous lines, originating from independent transformation events in T1 (second column) and T2 (in addition to those already found in T1) (third column). Column four: number of independently transformed lines with enhanced folate content. Column five: the maximal percentage of polyglutamylation per construct (if significantly enhanced).

Line	% Δ folate level month 2	% Δ folate level month 4
WT	37.00	10.25
GA9.15	-3.52	-8.30
GA26.5	66.41	63.75
GAmtF1.1.4.2	-27.95	-31.58
GAmtF1.1.11.4	-16.42	-25.03
GAmtF11.2.11.2	-42.88	-30.87
GAmtF15.3.3.1	126.38	73.81
GAmtF14.1.6.5	-51.01	-46.09
GAmtFsFBP32.1.2.2	-31.70	-21.63
2GAmtFsFBP14.1.5.3	-33.29	-10.26
2GAmtFCAFBP7.2.15	-53.82	-35.04
2GAmtFCAFBP10.1.16	-4.01	66.99
GAmtFGluB4FBP11.1.5	-15.37	33.86
GAmtFGluB4FBP44.1.14.14	-16.20	5.24
GAsFBP12.1.9.8	-20.97	-33.99
GAsFBP2.2.5.1	-18.93	-21.11
2GAsFBP3.1.6.8	-26.28	-37.24
GACAFBP3.1.3.2	-8.90	8.08
GACAFBP8.1.14.3	-21.10	-36.91
GACAFBP3.2.2	-3.87	-19.89
GAGluB4FBP17.2.4.6	-3.84	-12.41
GAGluB4FBP17.2.6.5	72.32	94.44
GActFsFBP 6.1.7.3	8.16	-17.30
GActFsFBP 6.1.8.2	1.36	-14.27
2GActFsFBP 4.2.21.3	4.58	5.42
2GActFsFBP 4.2.22.14	-54.70	-32.80
2GActFCAFBP 1.1.9.2	3.94	4.07

Supplemental table T3.2: Summary of change in folate content in the lines participating in the second stability experiment, expressed as percentages from zero time point folate level. Positive values indicate an increase in folate content, negative values a drop in folate level.

Line	Folate (µg/100g FW)	% poly	GTPCHI	ADCS	mtFPGS	ctFPGS	FBP
WT	15.07 ± 3.99	51.15	1	1	1	1	1
GA9	201.32 ± 48.60	15.36	588.8	1180	/	/	/
GA26	133.76 ± 28.57	15.23	161.4	4512	/	/	/
GAmtF1.1.4.2	792.89 ± 275.21	3.07	365.4	5748	59980	/	/
GAmtF1.1.11.4	600.83 ± 72.85	19.08	1367	47670	34150	/	/
GAmtF11.2.11.2	457.29 ± 81.78	32.78	320.2	12600	18480	/	/
GAmtF15.3.3.1	184.92 ± 58.03	30.23	74.07	1087	4671	/	/
GAmtF14.1.6.5	292.11 ± 122.37	15.70	602.2	2140	9285	/	/
GAmtFsFBP32.1.2.2	315.48 ± 35	18.35	142.4	2549	23690	/	99.22
2GAmtFsFBP14.1.5.3	175.14 ± 26.14	19.49	59.79	811.6	3773	/	1.1
2GAmtFCAFBP7.2.15	607.28 ± 83.92	22.93	1.6	5.372	21.2	/	12.59
2GAmtFCAFBP10.1.16	19.06 ± 6.38	29.70	/	/	/	/	/
GAmtFGluB4FBP11.1.5	140.74 ± 9.13	12.63	1.21	14.58	159	/	1.6
GAmtFGluB4FBP44.1.14.14	194.59 ± 25.15	22.66	324	2518	21300	/	10630
GAsFBP12.1.9.8	289.98 ± 86.45	9.40	108.1	214.7	/	/	972.7
GAsFBP2.2.5.1	264.95 ± 45.35	6.60	1941	10960	/	/	4858
2GAsFBP3.1.6.8	159.81 ± 48.66	14.19	442.7	3907	/	/	1225
GACAFBP3.1.3.2	548.63 ± 217.48	8.67	/	/	/	/	/
GACAFBP8.1.14.3	140.78 ± 26.34	22.63	196.4	1790	/	/	2.29
GACAFBP3.2.2	821.12 ± 168.04	3.81	72.44	197.2	/	/	173.5
GAGluB4FBP17.2.4.6	235.48 ± 21.26	17.85	391.9	1739	/	/	19.73
GAGluB4FBP17.2.6.5	245.61 ± 36.32	20.03	7.985	65.29	/	/	17.3
GActFsFBP 6.1.7.3	659.56 ± 106.47	11.47	1783	4911	/	2983	1254
GActFsFBP 6.1.8.2	497.33 ± 45.02	16.34	519.2	1353	/	2873	1088
2GActFsFBP 4.2.21.3	382.54 ± 67.55	18.61	523.8	11170	/	8466	341.7
2GActFsFBP 4.2.22.14	983.52 ± 382.08	5.81	345.3	2310	/	6412	196.2
2GActFCAFBP 1.1.9.2	1317.92 ± 367.65	6.38	6553	6639	/	25750	8285

Supplemental table S3.3: Overview of folate content, % polyglutamylation and normalized transgene expression of all lines participating in the second folate stability experiment. Folate content and polyglutamylation are averages from samples of month 2, 3 and 4, kept at -80°C; error bars indicate standard deviation. Expression analyses were performed by real-time quantitative PCR. Samples 2GAmtFCAFBP 10.1.16 and GACAFBP 3.1.3.2 got lost during RNA preparation. Rice tumor protein homologue (LOC_Os11g43900.1) and expressed protein (LOC_OS07g02340.1) were used as reference genes for normalization. Values are means of three technical replicates; error bars indicate standard error. Data analysis was performed using qBASE, based on the $2^{-\Delta\Delta Ct}$ method.

References

Blancquaert D, Storozhenko S, Loizeau K, De Steur H, De Brouwer V, Viaene J, Ravanel S, Rébeillé F, Lambert W, Van Der Straeten D. 2010. Folates and folic acid: from fundamental research toward sustainable health. Critical Reviews in Plant Sciences **29**, 14-35.

Blancquaert D, Storozhenko S, Van Daele J, Stove C, Visser R, Lambert W, Van Der Straeten D. 2013a. Enhancing pterin and para-aminobenzoate content is not sufficient to successfully biofortify potato tubers and *Arabidopsis thaliana* plants with folate. Journal of Experimental Botany **64**, 3899-3909.

Blancquaert D, Van Daele J, Storozhenko S, Stove C, Lambert W, Van Der Straeten D. 2013b. Rice folate enhancement through metabolic engineering has an impact on rice seed metabolism, but does not affect the expression of the endogenous folate biosynthesis genes. Plant Molecular Biology **83**, 329-349.

Choi SW, Friso S. 2005. Interactions between folate and aging for carcinogenesis. Clinical Chemistry and Laboratory Medicine **43**, 1151-1157.

De Brouwer V, Zhang GF, Storozhenko S, Van Der Straeten D, Lambert WE. 2007. pH instability of individual folates during critical sample preparation steps in prevision of the analysis of plant folates. Phytochemical Analysis **18**, 496-508.

De Brouwer V, Storozhenko S, Stove CP, Van Daele J, Van Der Straeten D, Lambert WE. 2010. Ultra-performance liquid chromatography-tandem mass spectrometry (UPLC-MS/MS) for the sensitive determination of folates in rice. Journal of Chromatography B **878**, 509-513.

Diaz de la Garza R, Quinlivan EP, Klaus SM, Basset GJ, Gregory JFIII, Hanson AD. 2004. Folate biofortification in tomatoes by engineering the pteridines branch of folate synthesis. Proceedings of the National Academy of Sciences of the United States of America **101**, 13720-13725.

Diaz de la Garza RI, Gregory JFIII, Hanson AD. 2007. Folate biofortification of tomato fruit. Proceedings of the National Academy of Sciences of the United States of America **104**, 4218-4222.

Endres M, Ahmadi M, Kruman I, Biniszkiewicz D, Meisel A, Gertz K. 2005. Folate deficiency increases postischemic brain injury. Stroke **36**, 321-325.

Hellemans J, Mortier G, De Paepe A, Speleman F, Vandesompele J. 2007. qBASE relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. Genome Biology **8**, R19.

Hossain T, Rosenberg I, Selhub J, Kishore G, Beachy R, Schubert K. 2004. Enhancement of folate in plants through metabolic engineering. Proceedings of the National Academy of Sciences of the United States of America **101**, 5158-5163.

Li GM, Presnell SR, Gu, LY. 2003. Folate deficiency, mismatch repair-dependent apoptosis, and human disease. Journal of Nutrional Biochemistry **14**, 568-575.

Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta Ct}$ method. Methods **25**, 402-408.

Ly A, Hoyt L, Crowell J, Kim Y-I. 2012. Folate and DNA methylation. Antioxidants and Redox Signaling **17**, 302-326.

Naqvi S, Zhu C, Farre G, Ramessar K, Bassie L, Breitenbach J, Perez Conesa D, Ros G, Sandmann G, Capell T, Christou P. 2009. Transgenic multivitamin corn through biofortification of endosperm with three vitamins representing three distinct metabolic pathways. Proceedings of the National Academy of Sciences of the United States of America **106**, 7762-7767.

National Institutes of Health – Office of Dietary Supplements. Available from: http://ods.od.nih.gov/factsheets/Folate-HealthProffesional/ Accessed August 2013.

Navarrete O, Van Daele J, Stove C, Lambert W, Van Der Straeten D, Storozhenko S. 2012. A folate independent role for cytosolic HPPK/DHPS upon stress in *Arabidopsis thaliana*. Phytochemistry **73**, 23-33.

Nunes ACS, Kalkmann DC, Aragão FJL. 2009. Folate biofortification of lettuce by expression of a codon optimized chicken GTP cyclohydrolase I gene. Transgenic Research **18**, 661-667.

Papakostas G, Cassiello CF, Iovieno N. 2012. Folates and S-adenosylmethionine for Major Depressive Disorder. Canadian Journal of Psychiatry **57**, 406-413.

Rueb S, Leneman M, Schilperoort RA, Hensgens L. 1994. Efficient plant-regeneration through somatic embryogenesis from callus induced on mature rice embryos (Oryza-sativa L). Plant Cell Tissue and Organ Culture **36**, 259-264.

Scarpella E, Rueb S, Boot KJM, Hoge JHC, Meijer AH. 2000. A role for the rice homeobox gene Oshox 1 in provascular cell fate commitment. Development **127**, 3655-3669.

Scott J, Rébeillé F, Fletcher J. 2000. Folic acid and folate: the feasibility for nutritional enhancement in plant foods. Journal of the Science of Food and Agriculture **80**, 795-824.

Scott JM, Weir DG. 1996. Homocysteine and cardiovascular disease. Quarterly Journal of Medicine **89**, 561-563.

Seshadri S, Beiser A, Selhub J, Jacques PF, Rosenberg IH, D'Agostino RB, Wilson PWF, Wolf PA. 2002. Plasma homocysteine as a risk factor for dementia and Alzheimer's disease. The New England Journal of Medicine **346**, 476-483.

Shane B. 1989. Folylpolyglutamate synthesis and role in the regulation of one-carbon metabolism. Vitamins and Hormones – Advances in Research and Applications **45**, 263-335.

Stanger O. 2004. The potential role of homocysteine in percutaneous coronary interventions (PCI): review of current evidence and plausibility of action. Cellular and Molecular Biology **50**, 953-988.

Storozhenko S, De Brouwer V, Volckaert M, Navarrete O, Blancquaert D, Zhang GF, Lambert W, Van Der Straeten D. 2007. Folate fortification of rice by metabolic engineering. Nature Biotechnology **25**, 1277-1279.

Yang L, Suzuki K, Hirose S, Wasaka Y, Takaiwa F. 2007. Development of transgenic rice seed accumulating a major Japanese cedar pollen allergen (Cryj 1) structurally disrupted for oral immunotherapy. Plant Biotechnology Journal **5**, 815-826.

Chapter 4

General conclusions and perspectives

General conclusions and perspectives

The first goal of this work was to investigate the effect of folate enhancement on general rice seed metabolism. In order to do so, samples were taken from developing rice seeds from day 8 after anthesis (DPA) up until day 28 DPA, when rice seeds were fully mature. Samples were taken from wild type (WT), a transgenic line overexpressing *Arabidopsis thaliana* (At) cDNA encoding GTP cyclohydrolase I (GTPCHI; a G line), transgenic rice overexpressing At cDNA encoding aminodeoxychorismate synthase (ADCS; an A line) and two rice lines overexpressing both cDNAs combined on a single T-DNA locus (GA lines; GA26.5 and GA9.15). These transgenic rice lines were previously created by Storozhenko and co-workers (Storozhenko et al., 2007). Folate and folate biosynthesis precursors *para*-aminobenzoate (*p*-ABA) and pterin levels were measured in all lines throughout rice seed development and transgene expression profiles were obtained. Based on these data, it was chosen to perform a transcriptomic profiling by microarray hybridization, where gene expression in GA lines was compared to that in WT at selected time points (8, 12 and 20 DPA) before and after folate levels were high. This way, we could investigate the effect of this folate accumulation on the metabolism of rice seeds. In total, 235 rice genes had an altered expression upon folate enhancement (Blancquaert et al., 2013a). In a similar study on folate biofortified tomato, the expression of only 14 genes was altered in GA tomato (Waller et al., 2010). Although gene coverage in this study was relatively small (the TOM2 array represents around 11,000 unigenes (Facella et al., 2008), whereas the tomato genome is predicted to have 34,727 protein-coding genes (The Tomato Genome Consortium, 2012)) as compared to the transcriptomic study in rice, it is clear that folate enhancement has a greater impact on rice seeds as compared to tomato fruit. Certainly, since the microarray experiment was conducted on developing rice seeds, more genes are detected with an altered expression, due to the fact that the 'experimental window' is broader than that used for tomato fruit (where only red-fruit stage was chosen for transcriptomic profiling). Interestingly, folate enhancement in rice did not affect the expression of the endogenous folate biosynthesis genes. In tomato fruit, 3 folate biosynthesis genes were affected by folate accumulation. This clearly indicates that fundamental differences on the regulation of folate metabolism exist between engineered rice seeds and tomato fruit and that the basal expression of downstream folate biosynthesis genes in biofortified rice seeds was sufficient to obtain high folate levels. In other words, if the expression of those (or one of those) three transgenes was not increased in engineered tomato, biofortification by engineering the *p*-ABA branch and the pterin branch would not have been sufficient to successfully biofortify tomato fruit with folate. Although a number of folate-related rice genes were revealed from which the expression was affected by folate enhancement, a lot of questions remain about their meaning. Moreover, a lot of genes with an unknown or a putative function showed up during this analysis and as long as their function remains unknown or unproven, no further conclusions can be drawn about their role in folate metabolism or secondary effects thereof. Furthermore, a number of genes involved in stress response and cell death showed an altered expression. Since *p*-ABA and pterins still accumulate in GA rice (and tomato), it is possible that the expression of some of these genes is affected by these folate biosynthesis precursors. Therefore, it is advisable to enhance the flux toward folate accumulation, by co-overexpressing downstream folate biosynthesis genes, such as mitochondrial HPPK/DHPS (an important regulatory enzyme, which catalyses the first two steps in the mitochondrial part of the folate

biosynthesis pathway) or FPGS, the last step in tetrahydrofolate polyglutamate biosynthesis (a 'pulling' strategy). This way, the accumulation of *p*-ABA and pterin could be decreased.

The second goal of this work was to biofortify potato tubers and *Arabidopsis thaliana* plants with folates. During this study it was convincingly proven that enhancing pterin and *p*-ABA levels (by metabolic engineering and/or *p*-ABA feeding) is not sufficient to enhance folate levels in those two species. It was concluded that another downstream 'bottleneck' in the flux toward THF accumulation is present in these two species and that most likely, co-overexpression of mitochondrial HPPK/DHPS (in combination with GTPCHI and ADCS overexpression) would be sufficient to biofortify potato tubers and Arabidopsis plants. Indeed, this bifunctional enzyme is predicted to be an important regulatory enzyme, since its DHPS domain is inhibited by dihydropteroate, dihydrofolate monoglutamate and tetrahydrofolate monoglutamate, which are intermediates of the folate biosynthesis pathway (Mouillon et al., 2002). Therefore, adapting engineering strategies to include overexpression of mitochondrial HPPK/DHPS , in combination with GTPCHI and ADCS, could not only be successful to improve folate content in recalcitrant crops, in which the two-gene strategy is inadequate, but also in other crops where a two-gene strategy works, to decrease the remaining high pterin and *p*-ABA levels, and further enhance folate content. The latter could be important, since little is known about possible toxicity of pterins (*p*-ABA is assumed to be harmless). Folylpolyglutamate synthetase (FPGS), an enzyme which lengthens the glutamate tail of tetrahydrofolate, could also be a good candidate for engineering. Since this protein catalyzes the last reaction in THF biosynthesis, its overexpression, in combination with GTPCHI and ADCS, could change engineering approaches from a 'pushing' to a 'pulling' strategy, hence forcing flux toward THF accumulation and possibly preventing pterin and *p*-ABA accumulation.
The question remains on which engineering strategy should be applied to successfully engineer folate levels in a staple crop. Clearly, further research is required to provide a satisfying answer, but it could be found by considering the following question: why was folate enhancement in tomato and rice successful with a two-gene approach while failing in potato and *Arabidopsis* (Blancquaert et al., 2013b)? It is clear that the activity of endogenous downstream steps was insufficient to guarantee THF accumulation in these two species. Whether this is caused by species-specific regulatory mechanisms of the biosynthesis pathway or by an insufficient activity of downstream enzymes, remains to be proven. A possible explanation can be found in species-specific and/or tissue-specific feedback mechanisms, compromising flux toward THF enhancement. Folate accumulation was targeted to endosperm (in rice) and pericarp (in tomato); tissues without the ability to differentiate into other tissues and whose primary function is to protect (endocarp) and supply the necessary nutrients and energy (pericarp and endosperm) to the embryos. On the other hand, potato tubers are derived from stem tissue (stolons) and besides being an important starch storage organ, it has the ability to redifferentiate and form new shoots; hence, it has a meristematic character. Since folate demand in meristematic plant tissue is high (due to its importance in C1 metabolism), it can be assumed that folate biosynthesis is tightly regulated in these cells, which could imply that engineering strategies should be adapted accordingly. A similar explanation for the unsuccessful attempts in *Arabidopsis* plants can be found in the importance of folates in photo-respiration and chlorophyll biosynthesis, where indeed C1 metabolism, and thus the demand of folate, should be tightly controlled in green tissues.

In the third part of this work, several approaches were followed to enhance folate stability in biofortified rice seeds. This was attempted by overexpression of two FPGS isoforms (mitochondrial and cytosolic FPGS) and by complexation with folate binding proteins. Although several transgenic rice lines were obtained with folate levels higher than ever reported in plants, especially in FBP overexpressing rice seeds, no clear results could be obtained for whether or not folate stability was improved in these transgenics. A final stability experiment is required to draw further conclusions on this topic.

References

Blancquaert D, Van Daele J, Storozhenko S, Stove C, Lambert W, Van Der Straeten D. 2013a. Rice folate enhancement through metabolic engineering has an impact on rice seed metabolism, but does not affect the expression of the endogenous folate biosynthesis genes. Plant Molecular Biology **83**, 329-349.

Blancquaert D, Storozhenko S, Van Daele J, Stove C, Visser R, Lambert W, Van Der Straeten D. 2013b. Enhancing pterin and para-aminobenzoate content is not sufficient to successfully biofortify potato tubers and Arabidopsis thaliana plants with folate. Journal of Experimental Botany **64**, 3899-3909.

Mouillon JM, Ravanel S, Douce R, Rébeillé F. 2002. Folate synthesis in higher plant mitochondria : coupling between the dihydropterin pyrophosphokinase and the dihydropteroate synthase activities. Biochemistry Journal **363**, 313-319.

Storozhenko S, De Brouwer V, Volckaert M, Navarrete O, Blancquaert D, Zhang GF, Lambert W, Van Der Straeten D. 2007. Folate fortification of rice by metabolic engineering. Nature Biotechnology **25**, 1277-1279.

The Tomato Genome Consortium 2012. The tomato genome sequence provides insights into fleshy fruit evolution. Nature **485**,635-641.

Waller JC, Akhtar TA, Lara-Nunez A, Gregory JFIII, McQuinn RP, Giovannoni JJ, Hanson AD. 2010. Developmental and feedforward control of the expression of folate biosynthesis genes in tomato fruit. Molecular Plant **3**, 66-77.

SUMMARY

Folates are important co-factors in one-carbon (C1) metabolism in all living organisms. They are synthesized *de novo* by plants and micro-organisms, hence, humans rely entirely on their diet as a sole folate source. Since most staple crops, such as potato and rice, contain low folate levels, folate deficiency is a global problem, which can cause an onset of disorders and diseases. In order to fight folate deficiency, a number of strategies, mostly involving synthetic folic acid, are currently implemented. Unfortunately, these approaches are often inaccessible for poorer populations. Therefore, enhancing folate levels in staple crops by metabolic engineering has been suggested as a complementary way to tackle folate deficiency worldwide. Half a decade ago, successful engineering attempts were reported on tomato and rice. The latter form the basis of this PhD dissertation. Due to the importance of folates in an onset of physiological processes, it was decided to investigate the effect of folate enhancement on general rice seed metabolism. Therefore, a transcriptomic study was conducted to see which rice genes had an altered expression due to folate accumulation. A total of 235 rice genes were differently expressed. Strikingly, no endogenous folate biosynthesis genes were revealed to have an altered expression upon folate enhancement.

Since potato is an important staple crop, containing low folate levels, the two-gene strategy, successfully implemented in tomato and rice, was used in attempts to biofortify potato tubers with folate. In parallel, this same approach was used to enhance folate content in *Arabidopsis* plants. Unfortunately, we have to conclude that this two-gene strategy is not sufficient to biofortify potato tubers and *Arabidopsis* plants with folate. This was confirmed by a *para*-aminobenzoate (*p*-ABA) feeding experiment on *Arabidopsis* seedlings engineered in pterin content. Although results on this part remain negative, it is of great importance for the scientific community and clearly shows that enhancing *p*-ABA and pterin content alone is not applicable to all plant species and tissues to obtain high folate levels. Most likely, another bottleneck downstream the folate biosynthesis pathway is present in potato tubers and *Arabidopsis* plants. These findings partially explain why after more than half a decade after the successful engineering attempts in tomato and rice, no other plants were successfully (> 10-fold increase as compared to control plants) biofortified with folate.

Folates are unstable molecules, susceptible to degradation. Hence, engineering folate levels in a crop alone is not sufficient to tackle the problem of folate deficiency. In our experiments, we clearly see that folate levels in engineered and wild type (WT) rice drop approximately 60% upon long storage (most folates were already degraded after 4 months) and that storage temperature (at least at non-freezing conditions) does not affect the rate of folate degradation. We tried to enhance folate stability in biofortified rice by increasing folate polyglutamylation (hence cellular retention and binding to folate dependent enzymes) and by complexation with folate binding proteins. Although the folate stability experiment on those lines needs to be repeated, we clearly see that overexpression of folate binding proteins (FBP), in combination with the necessary two genes for folate enhancement (GTP cyclohydrolase I (GTPCHI) and aminodeoxychorismate synthase (ADCS)) enables us to obtain transgenic

rice with folate levels up to approximately 2500 µg/ 100 g fresh weight (FW). These levels are the highest folate levels ever reported in plants and can be explained by a sequestration of folates by FBP.

SAMENVATTING

Folaten zijn belangrijke co-factoren in het een-koolstof (C1) metabolisme van alle levende organismen. Ze worden *de novo* aangemaakt door planten en micro-organismen. Bijgevolg zijn mensen volledig afhankelijk van hun dieet voor de inname van folaten. Aangezien de meeste voedingsgewassen, zoals aardappel en rijst, lage folaatgehalten bevatten, is folaatdeficientie een wereldwijd probleem, dat een hele reeks aan stoornissen en ziekten kan veroorzaken. Om folaatdeficientie tegen te gaan, worden een aantal strategieën, meestal via synthetisch foliumzuur, aangewend. Spijtig genoeg zijn deze strategieën meestal niet toegankelijk voor de armere populaties in de wereld. Als een complementaire aanpak om folaattekort te bestrijden, werd folaatbiofortificatie voorgesteld. Een half decennium geleden werden succesvolle *engineering* pogingen gepubliceerd in tomaat en rijst. De laatstgenoemde vormt de basis van dit doctoraatproject. Omwille van het belang van folaten in een reeks fysiologische processen, werd beslist om het effect van folaatverhoging op het algemene metabolisme van rijstzaden te onderzoeken. Om dit te doen, werd een studie op transcriptoomniveau uitgevoerd om na te gaan welke rijstgenen een differentiële expressie kennen door folaataccumulatie. In totaal werden 235 rijstgenen gevonden met een verhoogde of verlaagde expressie. Verrassend was het feit dat bij deze genen geen endogene rijstgenen aangetroffen werden die betrokken zijn in de folaatbiosynthese.

Aangezien aardappel een belangrijk voedingsgewas is, met lage folaateghaltes, werd aan de hand van de 2-genen strategie, die succesvol was in tomaat en rijst, getracht dit gewas te biofortifiëren met folaat. In parallel werd dezelfde aanpak gebruikt in *Arabidopsis* planten. Spijtig genoeg moeten we uit deze experimenten concluderen dat deze 2-genen strategie niet voldoende is om hogere folaatgehalten te verkrijgen in deze gewassen. Dit werd nog eens extra bevestigd door een *para*-aminobenzoaat (*p*-ABA) supplementatie experiment in *Arabidopsis* zaailingen, die verhoogde pterinegehalten hebben. Ook al zijn deze resultaten negatief, toch zijn ze van groot belang voor de wetenschappelijke wereld en werd duidelijk aangetoond dat het verhogen van *p*-ABA en pterinegehalten alleen niet toepasbaar is op alle planten en weefsels. Naar alle waarschijnlijkheid is er een andere limiterende stap aanwezig stroomafwaarts in de folaatbiosynthese pathway in aardappelknollen en *Arabidopsis* planten. Deze bevindingen kunnen deels verklaren waarom meer dan 5 jaar na de succesvolle pogingen in tomaat en rijst, er nog geen andere planten werden gerapporteerd met een verhoogd folaatgehalte (meer dan 10-voud).

Folaten zijn onstabiele moleculen, die vatbaar zijn voor degradatie. Bijgevolg is het verhogen van folaatgehalten in voedingsgewassen alleen niet voldoende om het probleem van folaatdeficiëntie aan te pakken. Uit onze experimenten blijkt dat folaatgehalten in wild type (WT) en gebiofortifieerde rijstzaden ongeveer 60% zakken bij lange bewaring (de meeste folaten waren reeds afgebroken na 4 maanden) en dat de bewaringstemperatuur (toch tenminste bij niet-vriezende condities) geen effect heeft op folaatdegradatie. We hebben getracht de folaatstabiliteit in gebiofortifieerde rijst te verhogen door verlenging van de glutamaatstaart te promoten (en aldus cellulaire retentie en binding door folaatafhankelijke enzymes te verhogen) en door complexatie met folaatbindende eiwitten. Ook al moet het folaatstabiliteitexperiment herhaald worden, toch zien we duidelijk dat bij een overexpressie van

folaatbindende proteïnen (FBP), in combinatie met de twee genen nodig om folaatverhoging te verkrijgen (GTP cyclohydrolase I (GTPCHI) en aminodeoxychorismaat synthase (ADCS)), er folaatgehalten verkregen worden tot ongeveer 2500 µg per 100 g versgewicht. Deze folaatgehalten zijn de hoogste ooit gerapporteerd in planten en kunnen verklaard worden door een sequestratie van folaten door FBP.

Addendum

Determination of the folate binding properties of sFBP, CAFBP and GluB4FBP

Determination of the folate binding properties of sFBP, CAFBP and GluB4FBP

One of the approaches to increase folate stability in biofortified rice comprises the use of folate binding proteins (FBP). Up until now, the existence of these proteins in plants remains unclear, but their occurrence, importance and action are well known in mammals (Henderson, 1990), where they form complexes with folates and protect them from degradation. A synthetic FBP, based on a mammalian FBP, with removal of the membrane anchorage signal peptide and codon-optimized for expression in plants, was synthesized for cloning. Since FBPs and folates bind in a 1-to-1 manner, massive accumulation of the former is required to maximize the pool of complexed folates in the folate enriched rice lines. Therefore, fusion sequences were created between the synthetic FBP and two abundant proteins (*Arabidopsis* β-carbonic anhydrase 2 and rice glutelin B4). Since it is possible that the fusion between FBP and CAβ2/GluB4 could compromise its folate binding capacities, the latter needs to be investigated.

Protein purification is needed in order to determine the folate binding capacities of the synthetic FBP and its fusions with glutelin B4 and beta-carbonic anhydrase 2. In a first attempt to purify the three FBP forms, *Escherichia coli* Rosetta-Gami B (DE3) pLysS was transformed with vectors pDESTsFBP17, pDESTCAFBP17 and pDESTGluB4FBP17 (Figure A1). These vectors contain an IPTG [1mM] inducible promoter to control FBP expression. However, the proteins could only be purified in denaturing conditions, since they resided in the insoluble fraction (Figure A2).

Therefore, it was decided to express and purify the proteins in *Pichia pastoris*. In order to do so, sFBP and its fusions where cloned into two different pPICz vectors: pPICZα (with an extracellular targeting sequence) and pPICZB (for intracellular expression). Each transgene will be expressed upon induction with methanol. A total of six constructs were made (pPICZα-sFBP, pPICZα-CAFBP, pPICZα-GluB4FBP, pPICZBsFBP, pPICZBCAFBP and pPICZBGluB4FBP). *Pichia pastoris* was transformed with these vectors and selected for colonies containing the transgenes. After isolation of the *Pichia* recombinants, the expression of the proteins was analyzed by SDS polyacrylamide gel electrophoresis. In the future, western blotting will be used to prove the expression of the proteins and to detect under what conditions the three proteins are optimally expressed. Subsequently, an up-scale expression experiment will be conducted and the sFBP, CAFBP and GluB4FBP proteins will be purified with the ProBond Purification System. Folate binding capacities of the purified proteins will be performed with the BiaCore system, based on surface plasmon resonance (SPR) (Leonard *et al*., 2011). This new technique enables the study of biomolecular interactions. In this respect, several folate species will be used as immobile ligands on the sensor chip surface and sFBP and its fusion proteins will be injected into the system. A change in SPR will be caused by the binding of these proteins to the ligands and affinity and chemical kinetics of these binding-interactions can be measured. This way, not only the binding of sFBP, CAFBP and GluB4FBP can be proven, also the preference of these proteins in binding the different folate species can be detected, a fundamental aspect which will provide new insights into the possibilities to stabilize folates in biofortified rice.

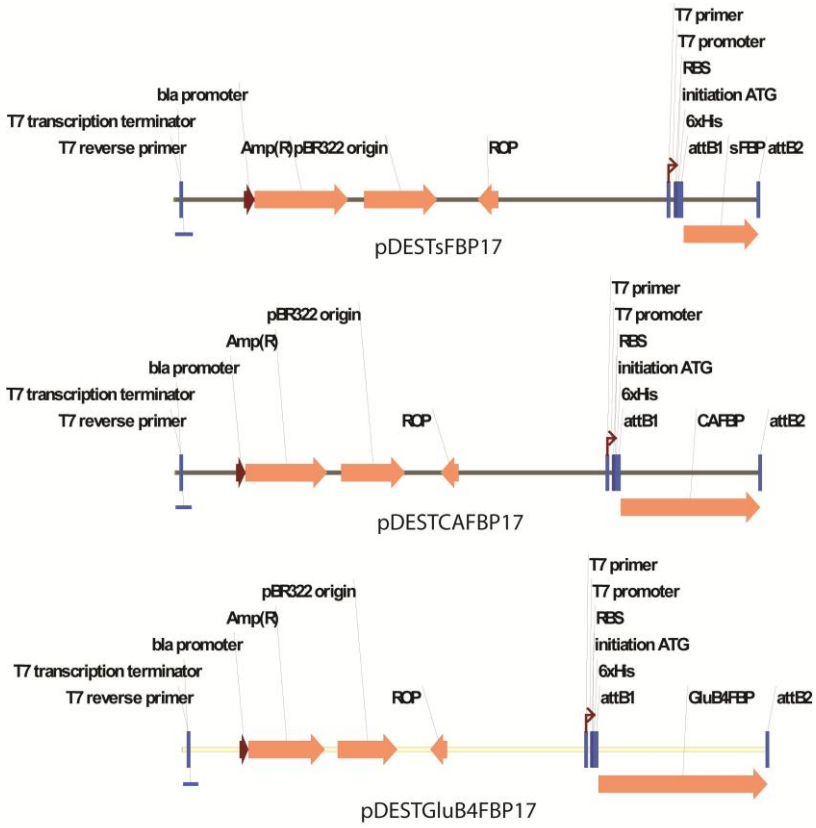


Figure A1. Schematic overview of the vectors used for sFBP, CAFBP and GluB4FBP expression and purification in E. coli.

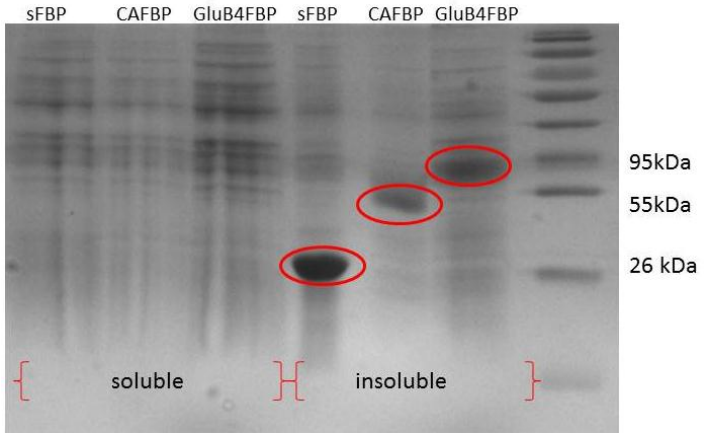


Figure A2. SDS polyacrylamide gel, stained with coomassie blue. The first three lanes comprise the soluble fraction, the latter three the insoluble fraction. All three FBP proteins could be detected in the insoluble fraction (red circles).

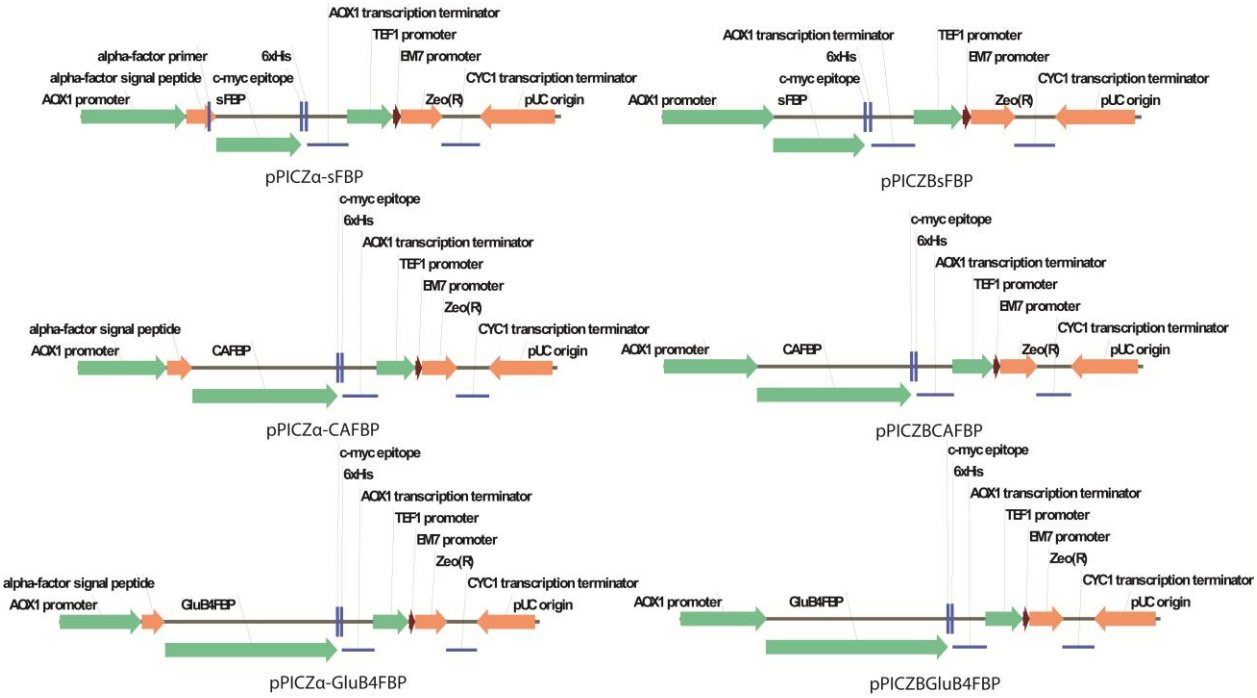


Figure A3. Schematic overview of the vectors used for sFBP, CAFBP and GluB4FBP expression and purification in *Pichia pastoris*.

References

Henderson GB. 1990. Folate-binding proteins. Annual Review of Nutrition **10**, 319-335.

Leonard P, Hearty S, O'Kennedy R. 2011. Measuring protein-protein interactions using BiaCore. In: Methods in Molecular Biology – Protein Chromatography: Methods and Protocols **681**, 403-418.

CURRICULUM VITAE

1. Personalia

Naam: Blancquaert

Voornamen: Dieter René Norma

Geslacht: Man

Nationaliteit: Belg

Burg. staat: Ongehuwd

Adres: Azaleastraat 17 bus 301, 9040 Sint-Amandsberg

Geboortedatum, -plaats: 11 maart 1984, Gent

GSM: 0485/81.70.98

E-mail: dieter.blancquaert@ugent.be

2. Opleiding

2.1 Voortgezette Academische Opleidingen

2008- 2013: Doctor in de Wetenschappen: Biotechnologie

Instelling: Universiteit Gent, Vakgroep Fysiologie, Laboratory of Functional Plant Biology

Titel proefschrift: Folate metabolism and biofortification in plants

Promotor: Prof. Dr. D. Van Der Straeten

2.2 Academische Opleidingen

2006-2007: 2e Licentie Biologie, optie Plantkunde (Universiteit Gent)

Behaalde graad: Grote onderscheiding

2005-2006: 1e Licentie Biologie, optie Plantkunde (Universiteit Gent)

Behaalde graad: Grote onderscheiding

2004-2005: 2e Kandidatuur Biologie (Universiteit Gent)

Behaalde graad: Onderscheiding

2003-2004: 1e Kandidatuur Biologie (Universiteit Gent)

Behaalde graad: Onderscheiding

2.3 Opleiding Secundair Onderwijs

2000-2006: secundair onderwijs (ASO), Latijn-Wiskunde (6uur)

Instelling: Sint-Janscollege, Sint-Amandsberg

3. Beroepsloopbaan

2008-2013: Predoctoraal onderzoeker, FWO-aspirant

Instelling: Universiteit Gent, Vakgroep Fysiologie, Laboratory of Functional Plant Biology (Prof.Dr. D. Van Der Straeten)

4. Onderwijservaring

2012-2013: Begeleiding Masterproef Pieter Dejonghe en Michiel Rydant

Begeleiding Bachelorproef Matthias Castelain, Juliette Roels en Heleen Vanhoolandt

Begeleiding Masterthesis Maarten Vandecauter

2011-2012: Begeleiding Masterproef Bruno Guillotin

Begeleiding Bachelorproef Noortje Gemoets, Nina Gorlé, Simon Stobbe en Jana Wyckmans

2008-2013: Onderrichten van de werkcolleges van het vak Plantenfysiologie (Prof. Dr. Ir. D. Van Der Straeten) aan 2^{e} Bachelor Biochemie/Biotechnologie studenten.

5. Wetenschappelijk Curriculum

A1:

Farré G, Blancquaert D, Capell T, Van Der Straeten D, Christou P, Zhu C. Engineering complex metabolic pathways in plants. Annual Review of Plant Biology (http://www.annualreviews.org/doi/abs/10.1146/annurev-arplant-050213-035825). Impact factor (2012): 26.654

Blancquaert D, Van Daele J, Storzhenko S, Stove C, Lambert W, Van Der Straeten D. 2013. Rice folate enhancement through metabolic engineering has an impact on rice seed metabolism, but does not affect the expression of the endogenous folate biosynthesis genes. Plant Molecular Biology **83**, 329-349. Impact factor (2012): 3.518

Blancquaert D, Storozhenko S, Van Daele J, Stove C, Visser RGF, Lambert W, Van Der Straeten D. 2013. Enhancing pterin and *para*-aminobenzoate content is not sufficient to successfully biofortify potato tubers and *Arabidopsis thaliana* plants with folate. Journal of Experimental Botany **64**, 3899-3909. Impact factor (2012): 5.242

De Steur H, Blancquaert D, Gellynck X, Lambert W, Van Der Straeten D. 2012. Ex-ante evaluation of biotechnology innovations: the case of folate biofortified rice in China. Current Pharmaceutical Biotechnology **13**, 2751-2760. Impact Factor (2012): 2.690

De Steur H, Gellynck X, Blancquaert D, Lambert W, Van Der Straeten D, Qaim M. 2012. Potential impact and cost-effectiveness of Multi-biofortified rice in China. New Biotechnology **29**, 432-442. Impact Factor (2011): 31.1

Blancquaert D, Storozhenko S, Loizeau K, De Steur H, De Brouwer V, Viaene J, Ravanel S, Rébeillé F, Lambert W, Van Der Straeten D. 2010. Folates and folic acid: from fundamental research toward sustainable health. Critical Reviews in Plant Science **29**, 14-35. Impact Factor (2010): 3.821

Storozhenko S, De Brouwer V, Volckaert M, Navarrete O, Blancquaert D, Zhang G-F, Lambert W, Van Der Straeten D. 2007. Folate fortification of rice by metabolic engineering. Nature Biotechnology **25**, 1277-1279. Impact Factor (2007): 22.848

B1:

Blancquaert D, Navarrete O, Storozhenko S, De Steur H, Van Daele J, Dong W, Lei C, Zhang C, Stove C, Gellynck X, Viaene J, Lambert W, Van Der Straeten D. 2012. Biofortified rice to fight folate deficiency. In: The Handbook of Food Fortification: From Concepts to Public Health Applications. Chapter 25. Humana Press.

De Steur h, Gellynck X, Blancquaert D, Storozhenko S, Liqun G, Lambert W, Van Der Straeten D, Viaene J. 2012. Market potential of folate biofortified rice in China. In: The Handbook of Food Fortification: From Concepts to Public Health Applications. Chapter 28. Humana Press.

6. Wetenschappelijke onderscheidingen en prijzen

04-07-2007: Prijs Pierre Verkerk, beste scriptie Biologie optie Plantkunde

Titel scriptie: Metabolische engineering van de biosynthese van folaten (vitamine B9) in rijst

7. Deelname aan Internationale Wetenschappelijke bijeenkomsten

1-2 december 2010: New Biotrends in green chemistry (Biotrends, Dortmund, Duitsland) (actieve bijdrage: postersessie)

6-7 november 2009: Food, sustainability and plant science – a global challenge (EMBL, Heidelberg, Duitsland)

13-14 november 2008: Bioactive compounds in plants – benefits and risks for man and animal (The Norwegian Academy of Science and Letters, Oslo, Noorwegen)

8. Overige verdiensten

2005-heden: Voorzitter Drosera VZW, Belgische Vereniging voor Liefhebbers van Vleesetende Planten

2010-heden: Lid Vakgroepraad, Vakgroep fysiologie, Faculteit Wetenschappen, Universiteit Gent

2005-2007: Lid Opleidingscommissie Biologie, Faculteit Wetenschappen, Universiteit Gent

9. Referentie

Prof. Dr. Dominique Van Der Straeten, Laboratory of Functional Plant Biology, Vakgroep Fysiologie, Universiteit Gent, E-mail: dominique.vanderstraeten@ugent.be, telefoonnummer: +32(0)9 264 5185, fax: +32(0)9 264 5333

Acknowledgements

Doing a PhD is rather a ‘lifestyle’ , not a ’job’. Trial and error, accompanied by frustrations and stress. Sometimes working on irregular moments, during weekends, in the evening, at night, early on in the morning, on holidays,… But just at that moment, when you get tired of all of this and reconsider to actually go and look for a job, little things, little experiments, all of a sudden do work. My first transformation vector. My first results. My first accepted publication. They gave me the strength to persue and to grow as a scientist, as an adult, as a human.

The folate project has always fascinated me and still does. I believe only a few PhD students are able to clarify their project to their family and friends, without losing them somewhere in the middle of their explanation (or at the very beginning in the worst case…).Well, I can. The social relevance of this project is higher than ever. Considering the fact that more and more negative outcomes are reported on folic acid fortification, there is a growing need for a valuable alternative. This is what we, all the members participating in this project, are trying to offer. And we succeed! Little by little, this project is expanding and more scientists get involved.

First, I would like to thank all jury members for actually reading this thesis, discovering mistakes, asking interesting questions and seeing things from another point-of-view! Although the closed defense was a very stressful moment for me, I learned a lot and some interesting ideas popped up during our discussion!

Second of all, I would like to thank all the members working in our group! Oana, Vera, Jolien, Katrien, Filip, Dajo, Pieter, Corina, Yuming, Sabine and Nuria. You all are great colleagues and I am very happy to see us growing as one group, inspiring each other and providing help were needed! Working in such a nice atmosphere is truly stimulating!

Third of all, a big thank you to all the other members, involved in our folate project! Oscar, Jeroen, Filip, Christophe, Willy, Hans, Xavier and Fabrice. It is a real pleasure working with all of you and your input, derived from your own expertise, has been invaluable to me and this thesis!

I would like to thank my friends and family and my parents in particular. Mama, papa, without you I would not be the person who I am now. Thank you for supporting me this much, especially now…

Last of all, I would like to express my gratitude to Dominique. You are a true leader, a source of inspiration and one of the strongest women I have ever met. You were always there for me, as a superior, as a scientist, as a friend. You gave me the strength to cope with difficult moments and situations. Thank you!