Towards nutritionally complete crops: the creation of FoliFerA potato

Author: Laura Virgínia Ferreras Custódio[1,2]
Supervisors: Prof. Dr. Dominique Van Der Straeten[1] and Prof. Dr. Isabel Sá-Correia[2]

1-Laboratory of Functional Plant Biology, Department of Biology, Ghent University, Belgium
2-Bioengineering Department, Instituto Superior Técnico, University of Lisbon, Portugal

Abstract

Micronutrient malnutrition, a condition named “The hidden hunger”, affects mainly pregnant and lactating women and children. Although clinical symptoms are not always evident, lifelong health complications arrive from this. Different health improvement programs have been established to tackle this problem. Biofortification programs have gained increasing attention since they represent a cost-effective and sustainable initiative to complement low-nutritional value diets, requiring little cultural adaptation and being easily introduced in regular diets. Different approaches for biofortification exist, however transgenic methods are the most promising as they are highly flexible and relatively rapid to obtain. In this work the expression vectors for the creation of multi-biofortified potato tubers were designed and constructed. This relied on the utilization of a novel cloning strategy, named Golden Braid, which allows a modular and iterative vector construction process. In this way, fifty-nine different vectors were produced, that can be used in the future for plant transformation, aiming for the enhancement of provitamin A (beta-carotene), iron and folate (vitamin B9). Additionally, in search of appropriate promoters to use in this project, the potential of a *Solanum tuberosum* promoter in driving root specific expression was revisited, showing that there is a need for updating DNA sequences repositories with newly curated information.

Key words: Micronutrient deficiency, Biofortification, Metabolic engineering, Folate, Vitamin A, Iron

Introduction

Micronutrient deficiencies (MND) affect approximately two billion people worldwide, being also known as “The Hidden Hunger”. This is a format of malnutrition that results from an insufficient intake, or absorption capacity, of one or more micronutrients. Micronutrients are a broad category of molecules, including vitamins (like vitamin A, vitamin B9 or vitamin B12) and minerals (such as zinc, iron and iodine). These compounds are key to optimal nutrition (Bailey et al., 2015), since they are known to participate in a number of biochemical processes, required for normal cellular functions.

Iron deficiency is the most widespread nutritional deficiency worldwide, and it is directly correlated with anaemia, a condition that afflicts nearly 1/3 of the human population (Murgia et al., 2012). This deficiency has also been linked with bad pregnancy outcomes and impaired cognitive development (Bastian et al., 2016; Lynch et al., 2018). Two of the most prevalent vitamin deficiencies are those of vitamin A and vitamin B9 (folate). Vitamin A deficiency (VAD) is one of the biggest health problems in developing countries, leading to high rates of infection and mortality (Oruch and Pryme, 2012; Shrivastava et al., 2014), it is also a common cause of xeropthalmia (Allen et al., 2006). Additionally, it may contribute to the development of iron-deficiency related disorders, such as anaemia (Michelazzo et al., 2013). Folate deficiency has been directly linked with neural tube defects (NTDs) during pregnancy and megaloblastic anaemia (Blancquaert et al., 2010).

The major risk groups for micronutrient deficiencies are pregnant and women of reproductive age, young children and the elderly, especially in low-income countries (Muthayya et al., 2013). Nevertheless, MNDs can affect human health at any life stage, impairing proper cognitive development and optimal performance (Von Grebmer et al., 2014). Although it is difficult to pinpoint the exact factors that lead to the appearance of an MND, low nutritional-value diets are a very common cause, particularly in low-income countries (Virot, 2007). These monotonous diets are mainly based on starchy crops, which deliver high caloric intakes,

however lack sufficient levels of minerals and vitamins. Because these populations do not possess sufficient access to micronutrient-dense foods, such as fruits and vegetables, they are commonly victims of deficiencies in multiple micronutrients (Brinch-Pedersen et al., 2007).

Biofortified foods have gained increasing attention for their potential role in tackling malnutrition, as this strategy aims for the generation of crops with enhanced micronutrient content, that could reach rural and poorer areas and would not depend on sophisticated delivery systems (Garcia-Casal et al., 2017).

Biofortification technology can be used to alter the micronutrient content of an existing crop through one of the following methods: agronomical interventions, conventional plant breeding or genetic engineering (Zhu et al., 2007). Despite the higher popularity and public acceptance of conventional breeding and agronomic interventions, recombinant DNA approaches present the greatest promise in terms of flexibility, given that they are not constrained by the host endogenous gene pool and allow organ-specific targeting (Mayer et al., 2008). Their aim is to modify endogenous metabolic pathways, or introduce new pathways, that ultimately result in an enhancement of the target product biosynthesis, lower concentrations of an unwanted molecule, or higher target product accumulation (Blancquaert et al., 2017).

The main staple crops, such as wheat, rice and potato, are generally insufficient providers of different micronutrients, and it is known that complicated and unstable socio-economical situations promote a higher consumption of these foods by large percentages of the population. To obtain plant crops that can address the full scope of MNDs, multi-biofortification appears to be the most prominent approach, since it could tackle multiple deficiencies at once. This will rely on multigene engineering strategies, that can be achieved through recombinant DNA technology (Naqvi et al., 2009).

Currently, food security is not a reality for a large proportion of the world population. Potato tubers hold the potential to help in the alleviation of these concerns, as they constitute a high value crop, that already is heavily consumed by many affected populations. Also, their natural nutritional value makes them an ideal vehicle, as little addition of specific micronutrients is necessary, in order to make it a complete crop. In this sense, investigation into meliorating its nutritional status is of outmost importance.

The main objective of this work is to develop the molecular tools needed for the multi-biofortification of *Solanum tuberosum* cv. Desireé with folate, β-carotene (provitamin A) and iron. This is to be achieved with special emphasis in the promoters and terminators that regulate the expression of the various chosen genes, in order to maximize expression in an organ-specific manner. The creation of these multi-gene expression vectors followed an iterative process dependent on restriction-ligation reactions, named Golden Braid 3.0 cloning (Vazquez-Vilar et al., 2017). This is a cloning technology that allows the exchange and assembly of DNA fragments in a modular framework, simplifying the assembly of complex expression vectors.

Materials and Methods

Vector construction

gDNA extractions and cDNA preparation

Genomic DNA was extracted from five different adult plants, namely, *N. benthamiana*, *S. tuberosum*, *S. lycopersicum*, *A. thaliana* and *P. sativum.* For all but the latter, for which only seeds were available, leaf tissue was collected and used for the extractions with the Invisorb Spin Plant Mini Kit (Invitek Molecular).

A cDNA sample from *A. thaliana* was prepared from extracted mRNA, obtained using the GeneJET plant RNA purification kit (Thermo Fischer Scientific). First, the extracted RNA was submitted to a DNase treatment, including Ribolock RNase inhibitor to minimize RNA degradation during the clean-up procedure (extended ThermoFischer Scientific RNA purification protocol). Second, a solution containing 1µg of the extracted mRNA was converted to cDNA through the iScript

cDNA synthesis kit (Bio-Rad), which contains random primers, in addition to the polydT oligonucleotides.

Restriction-Ligation Assembly Reactions

Restriction-ligation reactions were prepared as described in Sarrion-Perdigones et al., 2011, except that the DNA inserts were normalized to 60 fM. The insertion of assemblies / DNA fragments into the destination vectors relied on *BsaI* or *BsmBI* restriction enzymes (Thermo Fischer Scientific), as well as T4 ligase (Thermo Fischer Scientific). Each reaction mixture underwent 25 digestion/ligation cycles of 2 minutes at 37℃ and 5 minutes at 16℃, followed by three temperature steps, first, 37℃ for 5 minutes, second, 50℃ for 30 minutes, third, 80℃ for 10 minutes. The whole reaction was used to transform *Escherichia coli* TOP10 competent cells through heat-shock. The resulting bacterial suspension was plated in solid Luria-Bertani (LB) agar plates supplemented with 5-Bromo-4-chloro-3-indolyl-β-D-galactopyranoside acid (X-gal, 40 µg/mL) and isopropylthio-β-galactosidase (IPTG, 0.5 mM) allowing the white/blue selection of clones. Also, the solid media contained specific antibiotics, namely, chloramphenicol (50 µg/mL), kanamycin (50 µg/mL), or spectinomycin (100 µg/mL) for selecting transformed bacteria. White colonies were chosen for colony PCR, and the positive clones used to start liquid cultures of LB medium (37℃, 200 rpm, overnight) supplemented with the proper antibiotic. These were used for plasmid extraction using the GeneJET MiniPrep Plasmid extraction kit (Thermo Fischer Scientific). To assess if the assemblies were correct, the vectors were subjected to sequencing, restriction analysis or specific PCR reactions.

GBPart(s) Domestication

The DNA modules to be utilized in this cloning framework are named GBParts. These are obtained through PCR amplification with Phusion High-Fidelity DNA Polymerase (Thermo Fischer Scientific), from suitable templates (gDNA, cDNA or plasmids) and assembling them in UPD2 vectors. For this, specific GB primers were used (synthesized by IDT DNA), these contained a terminal region composed by a *BsmBI* recognition sequence, a cleavage site for cloning into UPD2 and a four-nucleotide barcode. Amplicons with correct size were purified using the GeneJET PCR product purification kit (Thermo Fischer Scientific) and quantified in a Nano Drop Spectrophotometer. After, they were used for a *BsmBI*-mediated restriction-ligation reaction (domestication). The resulting solution was used to transform *E. coli* TOP10 competent cells, which were plated in LB plates supplemented with 5-Bromo-4-chloro-3-indolyl-β-D-galactopyranoside acid (40 µg/mL) and isopropylthio-β-galactosidase (0.5 mM) and chloramphenicol (50 µg/mL). Positive white colonies were identified through colony PCR (T_a = 54℃; Extension time = 3 minutes) and targeted for plasmid extraction. Afterwards, the assemblies were sequenced (Macrogen) and their nucleotide sequences deposited in the Golden Braid database.

Domestication of C-terminal tagged coding sequences

To construct C-terminal tagged versions of the expression cassettes, the purified UPD2 fragments were used as a template for new amplification reactions, with the same forward primer as before and a new reverse primer containing in its tail, the coding sequence for different tags, before the STOP codon. For Coding Sequence 1 (CDS1), a single patch of 467 bp covering the initial 5' region was synthesized and used in a domestication reaction with a second patch of 3452 bp, which was obtained through PCR amplification of the remaining sequence for the synthetic β-carotene structure. Moreover, the peptide tag sequences were codon optimized for expression in *S. tuberosum*.

Domestication of recombination sites

The recombination sites were obtained through the annealing of two complementary oligonucleotides (synthetized by IDT DNA). These were subject to an annealing protocol, in which equimolar amounts of each oligo were mixed in annealing buffer (10mM Tris-HCL pH 8.0, 1mM EDTA pH 8.0, 50mM NaCl), and annealed using a program as follows: 5 minutes at 95 ºC; 1 minute at a range of temperatures, starting at 95 ºC and decreasing 1ºC until the melting temperature (77 ºC); 30

minutes at 77 ºC; 1 minute at a range of temperatures, starting at 77 ºC and decreasing 1 ºC until 25ºC. The obtained double-strand DNA fragment was directly used in a domestication reaction, as described previously.

Multipartite assembly

The second set of restriction-ligation reactions allows the assembly of multiple domesticated GBParts into a single expression cassette, named transcription unit (TU). The orderly assembly of the GBParts relies on the four-letter codes present in each fragment and is dependent on a *BsaI*-mediated restriction. The final assemblies were inserted in a pDGBα vector. Then, the reaction result was used to transform *E. coli* TOP10 competent cells, which were plated in LB plates supplemented with 5-Bromo-4-chloro-3-indolyl-β-D-galactopyranoside acid (40 µg/mL) and isopropylthio-β-galactosidase (0.5 mM) and kanamycin (50 µg/mL). White colonies were screened through colony PCR by the utilization of GBPart specific primers. Positive clones were targeted for plasmid extraction. Afterwards, the plasmids sequence was confirmed through restriction analysis (1h, at 37 ºC), and their nucleotide sequences deposited in the Golden Braid database.

Bipartite Assembly

The third set of restriction-ligation reactions aimed at combining different transcriptional units in a single expression vector. First, the purified TU-containing alpha vectors were subject to *BsmBI*-mediated restriction-ligation reactions. This allowed the assembly of two separate transcription units in a pDGBΩ vector. Then, the reaction result was used to transform *E. coli* TOP10 competent cells, which were plated in LB plates supplemented with 5-Bromo-4-chloro-3-indolyl-β-D-galactopyranoside acid (40 µg/mL) and isopropylthio-β-galactosidase (0.5 mM) and spectinomycin (100 µg/mL). White colonies were tested through colony PCR using GBPart specific primers and the positive clones were targeted for plasmid extraction. Afterwards, the plasmids sequence was confirmed through PCR analysis and restriction, and their nucleotide sequences deposited in the Golden Braid database. All subsequent reactions were of this type and followed the same logic of action.

Quantitative assessment of the expression level for a *Solanum tuberosum* promoter

Plants and growth conditions

Two *Solanum tuberosum* cv. Desiree representatives of the transgenic line TPPA (De Lepeleire et al., 2018) and two wild-type plants were obtained through vegetative propagation. These underwent a 2-week growth period *in vitro* in half-strength MS media (Murashige and Skoog), under 16/8 h light/dark daily cycles. Afterwards, the plantlets were transferred to soil and grown in day-long light cycles, at 24℃. Leaf and root tissue was collected after 6 weeks of growth.

RNA extraction and cDNA synthesis

RNA was extracted from leaf and root tissue of both wild-type and transgenic plants using the GeneJET plant RNA purification kit (Thermo Fischer Scientific). Afterwards, the solution was cleaned up with a DNase treatment, including Ribolock RNase inhibitor (Thermo Fischer Scientific) to minimize RNA degradation during the clean-up procedure. Then, a solution containing 1µg of the extracted mRNA was converted to cDNA through the iScript cDNA synthesis kit (Bio-Rad), which contains random primers, in addition the polydT oligonucleotides.

qPCR analysis

For the RT-qPCR, the KAPA Sybr fast qPCR kit (KAPA Biosystems) was utilized as a master reaction mix, following manufacturers recommendations. Each run consisted of a denaturation step at 95°C for 3 minutes, followed by 40 cycles of denaturation at 95°C for 15 seconds, annealing for 20 seconds, at the appropriate temperature for each primer pair, and extension at 72°C for 30 seconds. The collected data was analysed using the qBASE software based on the $2^{-\Delta\Delta}$Ct method (Hellemans et al., 2007; Livak and Schmittgen, 2001).

Results and Discussion

The utilization of the Golden Braid cloning framework simplified the construction of custom expression vectors

The Golden Braid technology consists in a modular DNA assembly framework, based off Golden Gate cloning. It can be divided into three separate phases, namely, domestications, multipartite assembly and bipartite assembly.

The first step consists on the domestication of the all the DNA sequences, that will be used in the expression vectors. This process relies in a PCR amplification using specific primers, containing a GB tail, that will allow the subsequent cloning of these fragments in an UPD2 vector. For some sequences, it is necessary to remove internal *BsmBI* and *BsaI* restriction sites, as these would interfere with subsequent cloning reactions. This is achieved through the PCR amplification of the target sequence in patches (small fragments), using specific primers that will introduce single mismatches in the restriction site, disrupting it. This was the case for CDS10 (2 patches), CDS7 (2 patches) and CDS3 (3 patches).

Four of the GBParts were synthetically produced (GenScript), to facilitate the cloning procedure, namely, CDS1 (as single a B3-B5 fragment); CDS8 (as a B4 fragment); CDS10 (as a B5 fragment); and T4 (as a B6-C1 fragment). Additionally, to ensure higher expression in tuber tissue and facilitate synthetic manufacturing, both the CDS1 and CDS10 were codon optimized for *S. tuberosum* expression using IDT codon optimization tool (https://eu.idtdna.com/CodonOpt).

Finally, four coding sequences were targeted for the addition of peptide tags in their 3´terminus, namely, CDS2, CDS3, CDS9 and CDS10. However, this was only performed after having all the desired fragments domesticated. As such, the “re-domestication” only required the PCR amplification of the sequence of interest from a purified UPD2 vector, with a modified reverse primer. For CDS1 there were difficulties in the amplification of the whole sequence, as such, an artificial patch was synthetically produced and used in a domestication reaction with a second patch, obtained through normal PCR amplification, with a modified reverse primer. A final list of all 24 domesticated fragments in this project is present below (table 1).

Table 1 – List of all domesticated fragments and their significance within the Golden Braid cloning technology. Abbreviations: TsP – tissue-specific promoter; IP – inducible promoter; T – terminator; CDS – coding sequence; TG – peptide tag; RS – recombination site.

ID	Function	Fragment Size (bp)
TsP1	**GBPart** Promoters (A1-B2)	642
TsP2		1525
TsP3		1122
IP		719
T1	**GBPart** Terminators (B6-C1)	966
T2		642
T3		441
T4		249
CDS1	**GBPart** Coding Sequence (B3-B5)	3795
CDS1_TG5		3819
CDS2		1234
CDS2_TG3		1264
CDS3		768
CDS3_TG4		795
CDS4		1077
CDS5		1032

ID	Funtion	Fragment Size (bp)
CDS6	**GBPart** Coding Sequence (B3)	2757
CDS7		1875
CDS8	**GBPart** Coding Sequence (B4)	72
CDS9	**GBPart** Coding Sequence (B5)	1398
CDS9_TG1		1416
CDS10		1545
CDS10_TG2		1569
RS	**TU (A1-C1)**	34

The correct size amplicons were purified. Then, they were used in the *BsmBI*-mediated domestication reactions, with a normalized concentration of 60 fM, to ensure equimolar ratios of the different patches (or fragments) and UPD2 vector. The resulting solution was used to transform competent *E. coli* TOP10 cells through heat-shock. After white colonies were screened through colony PCR, the positive clones were used for plasmid extraction. After sequencing pure vector samples, their sequences were deposited in the online Golden Braid database.

The application of the UPD2 vector as a universal entry vector for all domestications, allowed the utilization of the same DNA fragments in more than one cassette, without requiring any adjustments to the vector. Also, since the domestication reaction introduces the four-letter codes in the DNA fragments, the vector backbone remains unchanged. This allows greater flexibility when choosing the vectors for assembly, eliminating concerns regarding compatibility of the fragments with the vectors. Moreover, sequence-verifying all the domestication products eliminates, in theory, the need for future sequencing.

During this phase the main difficulties happened when utilizing the online software for analysing possible internal *BsaI*/*BsmBI* restriction sites. It occurred that once that the sequence modification suggested by the program led to an aminoacid substitution in the final protein. This was easily fixed by switching the problematic codon, by a synonymous counterpart, however it points to the fact that manual verification is necessary to ensure that no problematic modifications occur. Furthermore, a major flaw of the online domestication software is the impossibility of creating GBParts smaller than 100 bp, preventing the simple analysis of some interesting DNA motifs, regulation sequences or recombination sites. As was the case for the recombination site. Although this was manually fixed, it remains an issue for future projects.

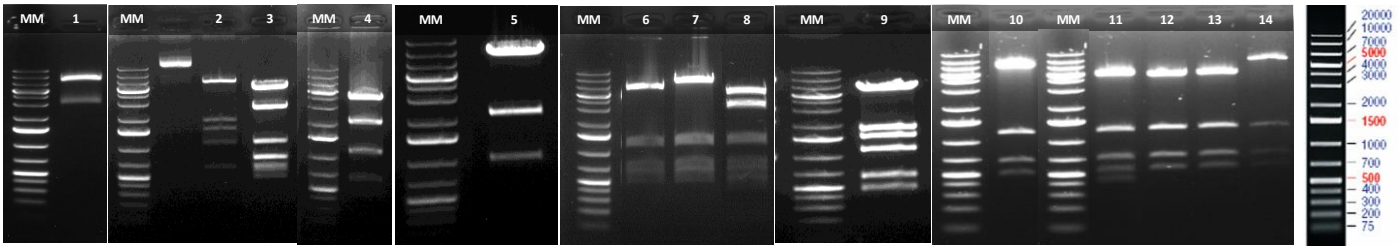
The second phase allows the composition of transcription units (multipartite assembly), which can be as simple, or as complex as desired, allowing the customization of the typical promoter::CDS::terminator structure accordingly. In fact, each TU can be composed of up to ten different components, that will all be all assembled in a single step.

The vectors containing the domesticated DNA parts were normalized to a 60 fM concentration and used for the transcription unit formation reaction, which is an identical program to the domestication, only difference being that it is dependent on a *BsaI* restriction-ligation reaction. A list of all the assembled transcription units is in table 2.

Table 2 – List of all the assembled transcription units. Abbreviations: TsP - tissue-specific promoter; IP – inducible promoter; T – terminator; CDS – coding sequence; RS – recombination site.

ID	Transcription Unit
TU_1	α1::TsP1_CDS1_T1
TU_1T	α1::TsP1_CDS1-Tag5_T1
TU_2	α2R::TsP2_CDS3_T3
TU_2T	α2R::TsP2_CDS3-Tag4_T3
TU_3	α1::TsP3_CDS2_T4
TU_3T	α1::TsP3_CDS2-Tag3_T4
TU_4	α1::TsP1_CDS6_CDS8_CDS9_T1
TU_4T	α1::TsP1_CDS6_CDS8_CDS9-Tag1_T1
TU_5	α2R::TsP2_CDS7_CDS8_CDS10_T2
TU_5T	α2R::TsP2_CDS7_CDS8_CDS10-Tag2_T2
TU_6	α2::CP_CDS4_T5
TU_7	α1::IP_CDS5_T5
TU_8	α1::RS
TU_9	α2::RS

After the previously mentioned reaction solution was used to transform competent *E. coli* TOP cells through heat-shock, positive white clones were identified through colony PCR. Then, we proceeded with plasmid extraction and ascertaining if the assemblies were correct through restriction analysis (figure 1).



Restriction Reactions with BamHI: 1 – TU_1 (8577, 3119, 87); **Restriction Reactions with Ava I:** 2 – TU_4 (6556, 1887, 1541, 1131, 630, 480); 3 – TU_5 (5090, 2597, 1131, 776, 773, 630, 480) ; 4 – TU_6 (4200, 2636, 1131, 630, 480); **Restriction Reactions with Hind III:** 5 – TU_5T (8961, 2138, 980); **Restriction Reactions with Ava I:** 6 – TU_3 (6303, 1131, 630, 480, 470) ; 7 – TU_2 (6873, 1131, 630, 480); 8 – TU_4T (6556, 1905, 1544, 1131, 630, 480); 9 – TU_1T (8163, 1403, 1131, 630, 480) ; 10 – TU_2T (6903, 1131, 630, 480); 11 – TU_3T (6303, 1131, 630, 480, 470) ; 12 – TU_8 (4168, 1131, 630, 480); 13 – TU_9 (4170, 1131, 630, 480) ; 14 – TU_7 (6383, 1131, 630, 480)

Figure 1 - Electrophoresis in an 1% agarose gel of the restriction reactions with the TU-containing alpha vectors. Each reaction was performed with 500 ng of pure plasmid, 1U/uL of restriction enzyme and 1X restriction buffer. The restrictions had a duration of 1 hour and were set at 37 ºC. Abbreviations: MM – Gene Ruler 1kb Plus DNA ladder (Thermo Fischer).

The restriction patterns for all the TU-containing alpha vectors matched the *in-silico* predictions (produced by the Vector NTI software), however, in reaction 9 (figure 1), there is a clear extra band above the expected 1.4 kb restriction fragment. This is possibly genomic DNA contamination that wasn't properly removed during the vector purification step, since all the other bands correspond specifically to the predicted restriction pattern. Specifically, the 1131, 630 and 480 bp bands obtained from backbone specific restriction, the 1403 bp band obtained from restriction between the 3' end of CDS1_Tag5 and 141 bp downstream of terminator 1 (T1), and the 8163 bp band obtained from backbone specific restriction and TU-specific restriction. After this confirmation step, it was time to proceed to next step in vector assembly.

The final phase consists in the pairwise combination of TUs in a third destination vector. This process was repeated several times to allow different TU combinations, in light of enhancing one, two or three micronutrients. These reactions allowed to create a set of expression cassettes with peptide tags (herein referred as tagged vectors), and a separate set without tags (herein referred as non-tagged vectors). The latter containing a self-excision recombination system module, that will permit controlled removal of the selection marker gene, after plant transformation. A simple overview of the several binary assemblies is present in table 3.

Because there was not an even number of transcription units, a stuffer vector was used to allow for pairwise combinations. The Golden Braid kit provides a collection of different level complementary stuffer vectors, which assist for bipartite assemblies. In this case, the stuffer vector contained a small non-coding intergenic sequence of *S. lycopersicum* DNA in place of the TU.

Table 3 – List of all the binary assemblies.

ID	Binary Assembly
Tagged Vectors	
1	Ω2::TU_4T::TU_5T
2	Ω2::TU_3T::TU_2T
3	Ω2::TU_1T_SF
4	Ω1R::TU_1T_SF
5	α2::KanR::TU_4T::TU_5T
6	α2::KanR::TU_3T::TU_2T
7	α1::SF::TU_3T::TU_2T
8	α2::KanR::TU_1T::SF
9	α1R::SF::TU_1T::_SF
10	Ω2R::SF::TU_3T::TU_2T::KanR::TU_4T::TU_5T
11	Ω2::SF::TU_1T::SF::KanR:: TU_4T::TU_5T
12	α2::SF::TU_1T::TU_4T::TU_5T::KanR::TU_3T::TU_2T
Non-Tagged Vectors	
I	Ω2::TU_4::TU_5
II	Ω1::TU_3::TU_2
III	Ω1R::TU_1::SF
IV	α1::TU_8::TU_6::TU_7::TU_9
V	α2:: TU_3::TU_2:: TU_4::TU_5
VI	α2::TU_1::SF::TU_4::TU_5
VII	Ω2::TU_8::TU_6::TU_7::TU_9::TU_3::TU_2:: TU_4::TU_5
VIII	α2::SF::TU_1::TU_8::TU_6::TU_7::TU_9::TU_3::TU_2::TU_4 ::TU_5

The produced assemblies were used to transform competent *E. coli* TOP10 cells through heat-shock. White positive clones were identified through colony PCR and targeted for plasmid extraction. Afterwards, all the vectors were evaluated through PCR analysis. This consisted on the amplification of specific regions within the utilized transcription units. Both sets of assembled vectors allowed the amplification of correct size fragments, correspondent to sections of some of the transcription units constructed. In addition, to confirm that the primer pairs utilized weren't amplifying unspecific regions of the purified vectors, or possible residual DNA fragments, a set of negative and positive control reactions were performed (figure 2).

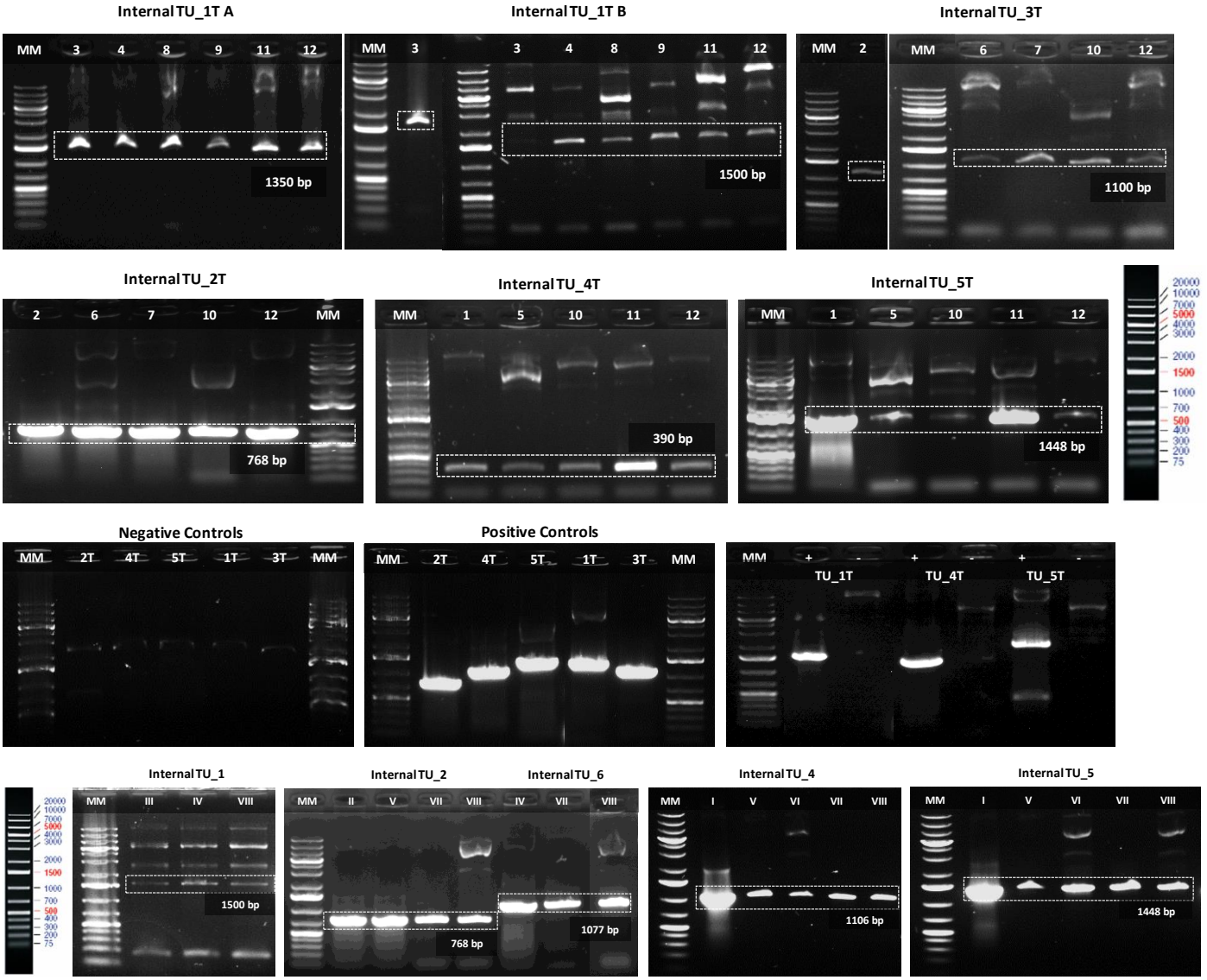


Figure 2 - Electrophoresis in a 1% agarose gel for the visualization of the PCR analysis for tagged vectors (TOP – 1 to 12) and non-tagged vectors (BOTTOM – I to VIII), as well as, positive and negative controls for the amplifications. Above each gel, a reference to transcription unit, whose internal region is targeted is indicated. Also, the expect size is indicated in each gel. Abbreviations: MM – Gene Ruler 1kb Plus DNA ladder (Thermo Fischer).

The positive controls templates consisted of purified and sequence verified UPD2 vectors, containing different domesticated fragments, or TU-containing alpha vectors, namely, TU_1, TU_4 and TU_5. The negative control template for all reactions was TU_3 vector, except the one concerning the TsP3 promoter amplifying primers, which had as template the TU_2 vector. All the reactions were performed equally to the PCR analysis, with the same number of cycles and identical annealing temperatures. This confirmed that all the primer pairs utilized for analysing the pure vectors were specific for their sequences. Also, they did not lead to the amplification of a similar size fragment in a vector that didn't contain the sequence of interest.

Despite the simplicity of the provided protocols, some technical difficulties emerged during the construction of these high complexity vectors. Whilst multipartite assemblies were easily obtained, bipartite assemblies were a lengthy process, with numerous attempts at each level of increased complexity. Perhaps due to the ever-increasing size of the produced vectors, which difficult the bacterial transformation step. Evidence of this was the presence of fewer colonies in each transformation step. Sequentially, there were some complications in the purification of vectors, which made the process of confirming if they were correctly assembled more difficult. Genomic DNA contaminations were very frequent, so the plasmid isolation protocol had to be optimized, to obtain very pure solutions.

Altogether, the GB framework is a simple to use cloning strategy, that had enough flexibility to allow the creation of all the envisioned vectors and permitted this to happen relatively quickly. In the end, 24 DNA fragments were domesticated, from genomic samples or pure plasmid preparations; 14 transcription units were assembled, and 20 binary assemblies were performed. This amounts to a total of 59 constructed vectors, going up to 30 kilobases, that will serve for the enhancement of one, two or three micronutrients in tuber tissue.

Assessing the potential of a candidate root-specific promoter from *S. tuberosum*

Most plant genetic engineering projects use constitutive promoter sequences, that help in gene overexpression, but sometimes at the cost of yield or physiological impairments (Lim *et al.,* 2012). In this sense, one of the main goals of this project was to utilize tissue-specific promoter sequences, that would drive a high transgene expression.

In search of good promoters, the capacity of a *S. tuberosum* promoter in driving high expression, within the root tissue, was assessed. The reason for investigating it derived from the abundance of research claiming that the gene family to which it belonged was expressed in root tissue (The Potato Genome Sequencing Consortium, 2011).

Two different sequences were identified and targeted for amplification from the *S. tuberosum* genomic DNA, however no amplificon was obtained after multiple attempts. Also, both the sequences lacked a characteristic 22 bp insert. This typical motif is present upstream of the transcription initiation position, and it was absent in both targeted sequences. A third similar promoter (Tissue-specific promoter 4.3 - TsP4.3), present in the laboratory was assessed, as it had been utilized in a previous project (DeLepeleire et al., 2018), although for tuber expression, not root. Once more, the characteristic motif was absent, however transgenic plants with genes under its control had already been produced, so a quantitative expression analysis was performed (figure 3). This preliminary test aimed for the comparison of expression, in different potato tissues, of one transgene under the control of this promoter and the endogenous gene expression, in both transgenic and wild-type plants.

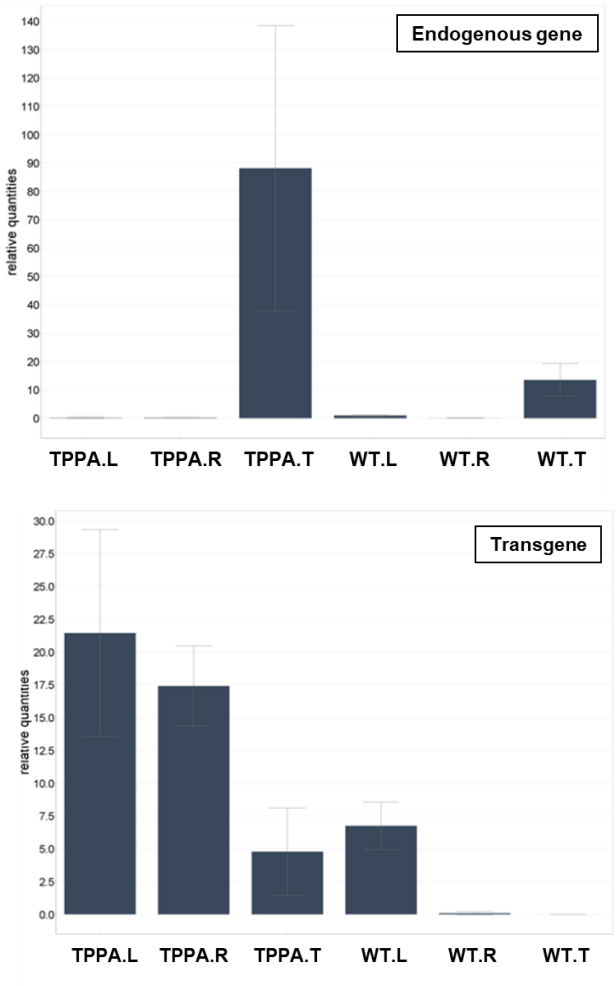


Figure 3 - Relative expression analysis of a transgene and endogenous gene in leaf, root and tuber tissue of *S. tuberosum* transgenic potato plant A (TPPA) and wild-type plants (value 1 represents WT mean relative expression for that tissue). Expression analysis was conducted using quantitative real-time PCR adult potato plants. Data analysis was performed using the qBASE software based on the 2–ΔΔCt calculation method. Normalization was achieved using *S. tuberosum* Elongation factor 1α and ribosomal protein L8 as reference genes (Nicot et al., 2005). The error bars represent standard deviation. Abbreviations: WT – S. tuberosum wild-type; TPPA – transgenic potato plant A; L – leaf tissue; R – root tissue; T – tuber tissue.

These results fail to replicate previous reports, also, the standard deviation in this assay was high, so little conclusions can be drawn without a more thorough analysis and more technical repeats. In transgenic plants (TPPA), endogenous gene expression was mainly identified in tuber tissue, with a very low, almost inexistent expression in root and leaf tissue. As for transgene expression, the higher expression in root and leaf tissue when comparing to the tuber, is clear. The wild-type plants recorded a much lower expression of

the endogenous patatin gene in tuber tissue, when comparing to the transgenic plants. Finally, as expected, the wild-type expression of the transgene was lower than the expression in transgenic plants.

Despite of its negative result, this primary assay was relevant in the sense that it shed a light on the need for updating expression data regarding regulatory sequences and genes, as well as, DNA sequences repositories.

Conclusion

Transgenic multi-biofortified crops could help alleviate micronutrient malnutrition, particularly in rural areas with little access to other alternatives. In this sense, creating effective molecular tools that allow the obtention of better plant food is imperative. The utilization of the Golden Braid cloning framework for the creation of all these expression vectors, facilitated the engineering process. It also allowed complete customization of the different transcription units, whilst providing simple to follow technical protocols.

In the future, the resulting transgenic potato tubers will be used to assess the value of this strategy in the enhancement of folate, provitamin A and iron concentrations, whilst assessing the potential micronutrient interactions and impact in plant physiology.

Acknowledgements

This document was written and made publicly available as an institutional academic requirement and as a part of the evaluation of the MSc thesis in Biotechnology of the author at Instituto Superior Técnico. The work described herein was performed at the Department of Biology of Ghent University (Ghent, Belgium), during the period February-September 2019, under the supervision of Prof. Dr. Dominique Van Der Straeten, and within the frame of the Erasmus programme. The thesis was co-supervised at Instituto Superior Técnico by Prof. Dr. Isabel Maria De Sá Correia Leite de Almeida.

References

Allen, L., World Health Organization, Food and Agriculture Organization of the United Nations, 2006. Guidelines on food fortification with micronutrients. World Health Organization; Food and Agriculture Organization of the United Nations, Geneva; Rome.

Bailey, R.L., West Jr., K.P., Black, R.E., 2015. The Epidemiology of Global Micronutrient Deficiencies. Ann. Nutr. Metab. 66, 22–33.

Bastian, T.W., von Hohenberg, W.C., Mickelson, D.J., Lanier, L.M., Georgieff, M.K., 2016. Iron Deficiency Impairs Developing Hippocampal Neuron Gene Expression, Energy Metabolism, and Dendrite Complexity. Dev. Neurosci. 38, 264–276.

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. Folate and Folic Acid: From Fundamental Research Toward Sustainable Health. Crit. Rev. Plant Sci. 29, 14–35.

Blancquaert, D., De Steur, H., Gellynck, X., Van Der Straeten, D., 2017. Metabolic engineering of micronutrients in crop plants: Metabolic engineering of micronutrients. Ann. N. Y. Acad. Sci. 1390,

Brinch-Pedersen, H., Borg, S., Tauris, B., Holm, P.B., 2007. Molecular genetic approaches to increasing mineral availability and vitamin content of cereals. J. Cereal Sci. 46, 308–326.

De Lepeleire, J., Strobbe, S., Verstraete, J., Blancquaert, D., Ambach, L., Visser, R.G.F., Stove, C., Van Der Straeten, D., 2018. Folate Biofortification of Potato by Tuber-Specific Expression of Four Folate Biosynthesis Genes. Mol. Plant 11, 175–188.

Garcia-Casal, M.N., Peña-Rosas, J.P., Giyose, B., consultation working groups, 2017. Staple crops biofortified with increased vitamins and minerals: considerations for a public health strategy: Micronutrient-biofortified crops in public health. Ann. N. Y. Acad. Sci.

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 Biol. 8, R19.

Lim, C.J., Lee, H.Y., Kim, W.B., Lee, B.-S., Kim, J., Ahmad, R., Kim, H.A., Yi, S.Y., Hur, C.-G., Kwon, S.-Y., 2012. Screening of tissue-specific genes and promoters in tomato by comparing genome wide expression profiles of Arabidopsis orthologues. Mol. Cells 34, 53–59.

Livak, K.J., Schmittgen, T.D., 2001. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2−ΔΔCT Method. Methods 25, 402–408.

Lynch, S., Pfeiffer, C.M., Georgieff, M.K., Brittenham, G., Fairweather-Tait, S., Hurrell, R.F., McArdle, H.J., Raiten, D.J., 2018. Biomarkers of Nutrition for Development (BOND)—Iron Review. J. Nutr. 148, 1001S-1067S.

Mayer, J.E., Pfeiffer, W.H., Beyer, P., 2008. Biofortified crops to alleviate micronutrient malnutrition. Curr. Opin. Plant Biol. 11, 166–170.

Michelazzo, F., Oliveira, J., Stefanello, J., Luzia, L., Rondó, P., 2013. The Influence of Vitamin A Supplementation on Iron Status. Nutrients 5, 4399–4413.

Murgia, I., Arosio, P., Tarantino, D., Soave, C., 2012. Biofortification for combating 'hidden hunger' for iron. Trends Plant Sci. 17, 47–55.

Muthayya, S., Rah, J.H., Sugimoto, J.D., Roos, F.F., Kraemer, K., Black, R.E., 2013. The Global Hidden Hunger Indices and Maps: An Advocacy Tool for Action. PLoS ONE 8, e67860.

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. Proc. Natl. Acad. Sci. 106, 7762–7767.

Nicot, N., Hausman, J.-F., Hoffmann, L., Evers, D., 2005. Housekeeping gene selection for real-time RT-PCR normalization in potato during biotic and abiotic stress. J. Exp. Bot. 56, 2907–2914.

Oruch, R., Pryme, I.F., 2012. The biological significance of vitamin A in humans: A review of nutritional aspects and clinical considerations 14.

Sarrion-Perdigones A, Falconi EE, Zandalinas SI, Juárez P, Fernández-del-Carmen A, Granell A, et al. (2011) GoldenBraid: An Iterative Cloning System for Standardized Assembly of Reusable Genetic Modules. PLoS ONE 6(7): e21622.

Shrivastava, S., Shrivastava, P., Ramasamy, J., 2014. Vitamin-A deficiency: A global cause of public health concern. Int. Med. J. Sifa Univ. 1, 29.

Stupar, R.M., Beaubien, K.A., Jin, W., Song, J., Lee, M.-K., Wu, C., Zhang, H.-B., Han, B., Jiang, J., 2006. Structural Diversity and Differential Transcription of

the Patatin Multicopy Gene Family During Potato Tuber Development. Genetics 172, 1263–1275.

The Potato Genome Sequencing Consortium. (2011). Genome sequence and analysis of the tuber crop potato. Nature, 475(7355), 189-195.

Virot, P., 2007. Preventing and controlling micronutrient deficiencies in populations affected by an emergency. World Health Organ. 2.

Von Grebmer, K., Saltzman, A., Birol, E., Wiesmann, D., Prasai, N., Yin, S., Yohannes, Y., Menon, P., Thompson, J., Sonntag, A., 2014. 2014 Global Hunger Index: The challenge of hidden hunger. Welthungerhilfe, International Food Policy Research Institute (IFPRI).

Zhu, C., Naqvi, S., Gomez-Galera, S., Pelacho, A.M., Capell, T., Christou, P., 2007. Transgenic strategies for the nutritional enhancement of plants. Trends Plant Sci. 12, 548–555.