



OPEN

Development of screening methods for functional characterization of UGTs from *Stevia rebaudiana*

Eva Petit¹, Monique Berger^{1✉}, Laurent Camborde², Veronica Vallejo³, Jean Daydé¹ & Alban Jacques¹

Glycosylation is a key modification that contributes to determine bioactivity and bioavailability of plant natural products, including that of terpenoids and steviol glycosides (SVglys). It is mediated by uridine-diphosphate glycosyltransferases (UGTs), that achieve their activity by transferring sugars on small molecules. Thus, the diversity of SVglys is due to the number, the position and the nature of glycosylations on the hydroxyl groups in C-13 and C-19 of steviol. Despite the intense sweetener property of SVglys and the numerous studies conducted, the SVglys biosynthetic pathway remains largely unknown. More than 60 SVglys and 68 putative UGTs have been identified in *Stevia rebaudiana*. This study aims to provide methods to characterize UGTs putatively involved in SVglys biosynthesis. After agroinfiltration-based transient gene expression in *Nicotiana benthamiana*, functionality of the recombinant UGT can be tested simply and directly in plants expressing it or from a crude extract. The combined use of binary vectors from pGWBs series to produce expression vectors containing the stevia's UGT, enables functionality testing with many substrates as well as other applications for further analysis, including subcellular localization.

Steviol glycosides (SVglys) are *ent*-kaurenoic diterpenoids produced and accumulated in *Stevia rebaudiana* leaves up to 25% of the dry matter¹. Biosynthesis of SVglys starts with the formation of geranyl-geranyl diphosphate (GGDP) in the chloroplast from the methyl-erythritol-4-phosphate (MEP) pathway²⁻⁴. GGDP then undergoes two cyclization steps, catalyzed by terpene cyclases to give *ent*-kaurene⁵. *Ent*-kaurene is subsequently oxidized into kaurenoic acid (KA) by kaurene oxidase (KO) in the endoplasmic reticulum (ER). Kaurenoic acid is the last step of a biosynthetic pathway shared by gibberellins and SVglys. Oxidation of KA leads to gibberellins, while its hydroxylation produces steviol that can be glycosylated on the two hydroxyl groups of its C-13 and C-19 to give SVglys. Glycosylation is catalyzed by UDP (uridine diphosphate)-glycosyltransferases (UGTs). To form steviol glycosides, UGTs transfer an activated sugar of a sugar-donor substrate on hydroxyl groups of an acceptor substrate. Glycosylation is a key and widespread modification of secondary metabolites, playing a pivotal role in the huge diversity of plant natural products⁶. Thus, more than 60 SVglys have been identified in *S. rebaudiana* leaves to date¹. Main SVglys are stevioside (ST), rebaudioside A, C, D, E, F (Reb A-F), dulcoside A (Dulc A) and steviolbioside. In *S. rebaudiana*, a wide phenotypic variability in SVglys profiles can be observed⁷⁻⁹. For example, some genotypes synthesize very high proportions of Reb A and/or Reb C, whereas others accumulate preferentially their corresponding substrates ST and Dulc A, respectively^{10,11}. These differences could be due to variability of the enzyme responsible for Reb A and Reb C biosynthesis, UGT76G1^{10,11}.

The first part of the SVglys biosynthesis leading to steviol formation is well known and characterized^{12,13}. The latter part, leading to SVglys from the steviol aglycone remains largely unknown as *S. rebaudiana* UGTs are poorly characterized, preventing a better understanding of the SVglys pathway. Nevertheless, several UGT genes of *S. rebaudiana* are available. For example, 12 candidates have previously been identified¹⁴. In a more recent study, Li et al. isolated 144 sequences suspected to be UGTs by re-analyzing the public transcript data of *S. rebaudiana* from NCBI (SRR1576548), among them 68 were unambiguously identified as UGT genes¹⁵.

When a UGT gene is available, one of the most common methods to identify its functionality is in vitro expression in a heterologous system¹⁶. Substrate specificity of plant UGTs involves recognition of a sugar acceptor as

¹Equipe Physiologie, Pathologie et Génétique Végétales (PPGV), INP-PURPAN, Université de Toulouse, 75 voie du TOEC, BP 57611, 31076 Toulouse Cedex 03, France. ²Laboratoire de Recherche en Sciences Végétales (LRSV), CNRS, Université Paul Sabatier (UPS), Toulouse, France. ³PepsiCo, 700 Anderson Hill Rd., Purchase, NY 10577, USA. ✉email: monique.berger@purpan.fr

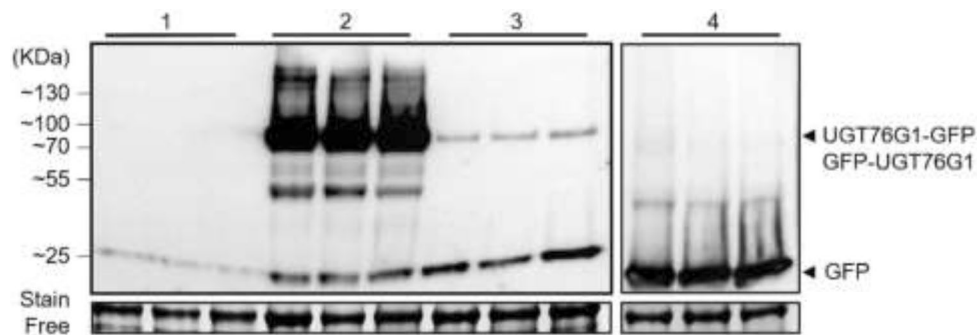


Figure 1. Anti-GFP western blot analysis of recombinant proteins expressed in *N. benthamiana* at 5 days post-infiltration (dpi). Recombinant proteins are shown with arrows. The figure presents 3 biological repetitions per condition. 1: non-transformed plants; 2: p35S::UGT76G1-GFP; 3: p35S::GFP-UGT76G1; 4: p35S::GFP.

well as a UDP-sugar donor¹⁷. In most instances, UGTs show a strong specificity for the UDP-sugar; even those exhibiting activities toward several sugar donors, often have a better affinity for only one of them¹⁷. UGTs show broad differences in their individual number of possible acceptors. Some UGTs are highly specific and only glycosylate few acceptors. For example, UGT73A14 of *Gentiana triflora* specifically glycosylates the 3'-hydroxy group of delphinidin-type anthocyanins having glucose at 3 and 5 positions¹⁸. Many other UGTs are regiospecific, thus they can glycosylate a large array of compounds, often part of the same metabolic family. Possible acceptor substrates tested in vitro can also be different. This is the case of UGT76G1 of *S. rebaudiana*, which glycosylates alcohols, substituted monophenols, polyphenols and glycosides¹⁹. Identification of an enzyme for a single glycosylation step using in vitro tests is limited by the panel of substrates used. For example, UGT74G1 and UGT73E1 appear to be steviol C-4 glucosyltransferases giving the 19-*O*-glucoside in *S. rebaudiana*^{14,15}. However, they were also found to glycosylate the same site to produce rubusoside from steviolmonoside. Additionally, these tests may disregard the substrate synergy present in plant or the affinity of enzymes for these different substrates. These ambiguities underpin the need to test the specificity of plant UGTs in vivo, to better understand the complex or uncommon metabolic pathways as well as the plasticity of metabolism in response to biotic and abiotic stresses.

A method to reveal the functionality of a plant UGT which takes into account their polyspecific nature, is the creation of mutants of which the phenotype is compared to the control plants. Two different methods are available to create genetically transformed plants: stable or transient transformations. Stable transformation consists in integrating a foreign DNA sequence in the genome of a plant of interest in order to modulate plant traits such as stress tolerance, disease resistance or metabolite production. In the opposite, transient transformation is a process leading to the introduction of a nucleic sequence with temporary expression and tissue specific localization. Transient transformation has several advantages compared to stable one, as it does not need to create stable lines.

This study deals with two simple and original methods to characterize functionality of stevia's UGTs after transient expression in *Nicotiana benthamiana* with the caveat that this system is not intended to characterize enzyme kinetics. These methods (1) avoid the purification step generated by recombinant protein expression in micro-organisms or plants, and allow to (2) quickly screen the functionality of an enzyme with individual substrates, (3) test a mix of proven acceptors, which might give clues on the substrate affinity of the enzyme in vivo.

Results

Recombinant proteins are expressed in *Nicotiana benthamiana*. To characterize UGT76G1, GFP was added in N or C terminus of the protein as tag may interfere with protein localization or activity^{20,21}. Both constructs p35S::UGT76G1-GFP and p35S::GFP-UGT76G1 were agroinfiltrated in *N. benthamiana* leaves, using *A. tumefaciens*. At 5 days post-infiltration (dpi), specific bands are observed on western blot at around 77 kDa, the expected size of UGT76G1-GFP and GFP-UGT76G1 proteins (Fig. 1). Strong bands corresponding to UGT76G1-GFP confirm the accumulation of the protein while weak bands are detected for the N-terminally tagged version (see Supplementary Fig. S1 for longer acquisition signal). This indicates that although both versions of UGT76G1 recombinant proteins are expressed, it seems that GFP N-terminal tag may interfere with the accumulation of the protein. Bands of great intensity at around 25 kDa corresponds to expression of GFP as control.

UGT recombinant proteins harbor a nucleo-cytoplasmic localization. Expression of recombinant proteins was verified by confocal laser scanning microscopy (CLSM) at 5 dpi. The GFP fluorescence was observed in leaves transformed with GFP, GFP-UGT76G1 and UGT76G1-GFP, confirming their expression (Fig. 2).

Recombinant GFP is visible in the cytosol for all three constructs due to the specific green fluorescence of filaments passing through the cell. Part of the specific fluorescence is observed in the endoplasmic reticulum, which appears as a diffuse green "grid" in the cell. This is probably due to the presence of recombinant proteins at the level of the maturation zone.

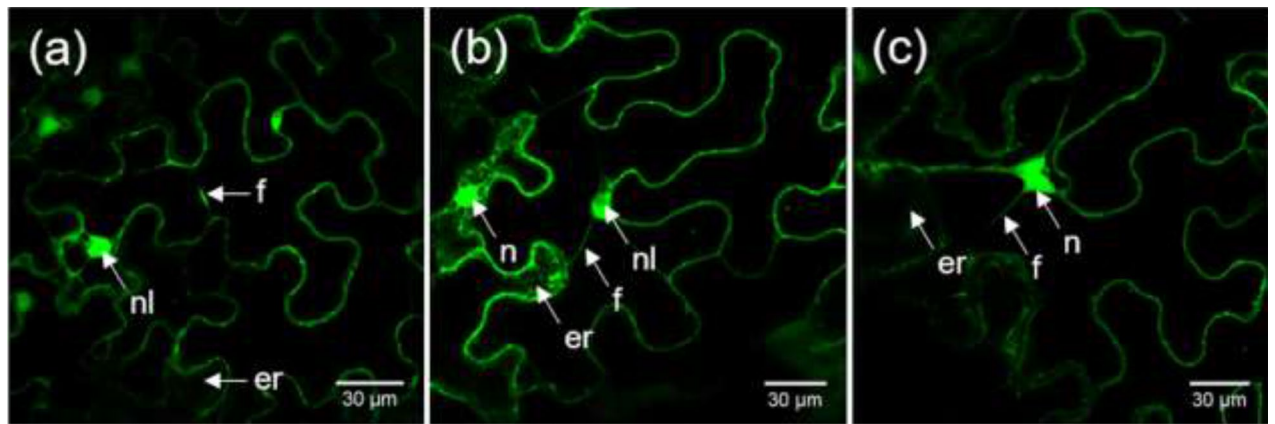


Figure 2. CLSM observations of recombinant proteins at 5 dpi (a) GFP, (b) GFP-UGT76G1, (c) UGT76G1-GFP. *er* endoplasmic reticulum, *f* cytosolic filaments, *n* nucleus, *nl* nucleolus.

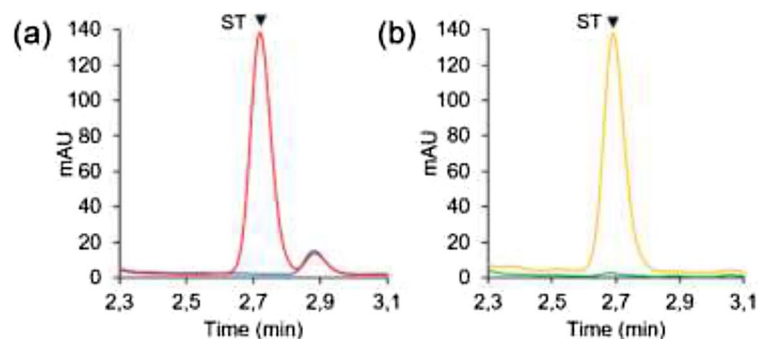


Figure 3. Glycosylation activity of non-transformed *N. benthamiana*. (a) In vitro test: a crude extract was produced from leaves of non-transformed *N. benthamiana*, and was analyzed by UPLC after a 2 h-incubation with (red) or without (blue) ST, (b) In planta test: leaves of non-transformed *N. benthamiana* were analyzed by UPLC, 24 h after infiltration of stevioside in leaves (yellow) or no infiltration (green). *ST* stevioside.

GFP fluorescence is also detected in the nucleus for all three constructs at 5 dpi (but not in the nucleolus). This localization could be due to passive diffusion of GFP. Some studies support the hypothesis of bidirectional diffusion of free GFP through nuclear pores, due to its low molecular weight^{22,23}. GFP-UGT76G1 and UGT76G1-GFP being bigger, one hypothesis to explain their nuclear fluorescence could be the degradation of recombinant proteins allowing fragments of free GFP to diffuse into the nucleus. This hypothesis is supported by our western blot analyzes at 5 dpi, where specific GFP band (26.4 kDa) was revealed in the corresponding lanes (Fig. 1).

***N. benthamiana* does not convert stevioside in steviol glycoside.** Stevioside (ST) and other SVglys were identified by their elution times and UV spectra using corresponding standards (see Supplementary Fig. S2). Chromatograms corresponding to leaves from non-transformed plants and that have been in presence of ST show a characteristic peak at around 2.7 min corresponding to ST (red and yellow curves), contrary to the negative control (blue and green curves) (Fig. 3). No new peak appears close to that of ST, meaning that non-transformed *N. benthamiana* does not have a UGT able to transform ST into Reb A. This result was confirmed by both functionality tests and allowed the continuation of activity trials on plants expressing UGT76G1.

UGT76G1 expressed in *N. benthamiana* produces rebaudioside A from stevioside. In vitro and in planta activity tests were then conducted on *N. benthamiana* leaves expressing UGT76G1. According to the chromatograms, leaves expressing UGT76G1 and that have not been in the presence of ST (blue and green curves), cannot synthesize Reb A naturally as no peak specific to SVglys appears (Fig. 4). In the presence of ST (red and orange curves), the UGT76G1 expressed in *N. benthamiana* catalyzes β -1,3 glycosylation to produce Reb A, whose characteristic peak appears on the chromatograms at 2.6 min.

GFP tag does not affect UGT76G1 functionality. Functionality of fusion proteins UGT76G1-GFP and GFP-UGT76G1 was then tested using the two methods. According to the chromatograms, the negative controls (corresponding to blue and green curves) do not synthesize Reb A as no specific peak appears (Fig. 5). In the presence of ST (red and orange curves), UGT76G1-GFP and GFP-UGT76G1 proteins expressed in *N. bentha-*

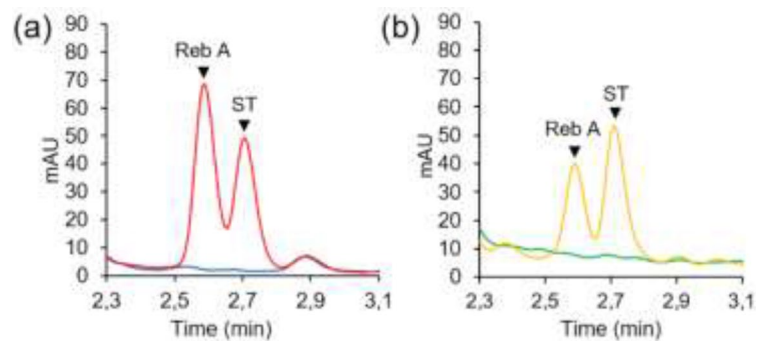


Figure 4. Activity of UGT76G1 expressed in *N. benthamiana* at 5 dpi. **(a)** In vitro test: a crude extract was produced from leaves of transformed *N. benthamiana* and was analyzed by UPLC after a 2 h-incubation with (red) or without (blue) ST, **(b)** In planta test: leaves of transformed *N. benthamiana* were analyzed by UPLC, 24 h after infiltration of ST in leaves (yellow) or no infiltration (green). *Reb A* rebaudioside A, *ST* stevioside.

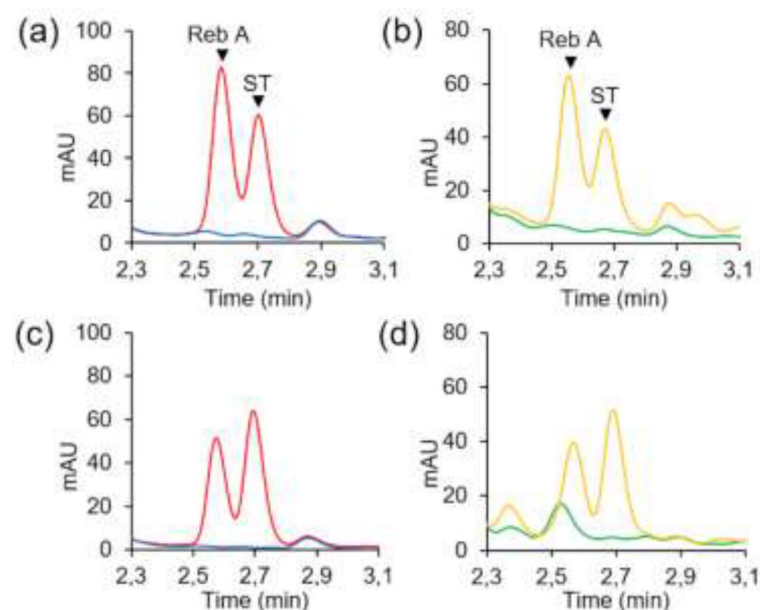


Figure 5. Activity of UGT76G1-GFP and GFP-UGT76G1 expressed in *N. benthamiana* at 5 dpi. **(a, c)** In vitro test, after a 2 h-incubation of a leaf extract with (red) or without (blue) ST, **(b, d)** In planta test, 24 h after infiltration of ST in leaves (yellow) or not (green). **(a, b)** p35S::UGT76G1-GFP, **(c, d)** p35S::GFP-UGT76G1. *Reb A* rebaudioside A, *ST* stevioside. Original chromatogram for **(a)** (red) is given in Supplementary Fig. S4 and compared with SVglys standards for peak identification.

miana catalyze the β -1,3 glucosylation to produce Reb A. The same results have been observed for UGT76G1 coupled with a polyhistidine tag UGT76G1-6xHis and 6xHis-UGT76G1 (see Supplementary Fig. S3).

These two methods of functional characterization can therefore be used for UGTs that intended to be purified or to be localized at the cellular level, provided that the conformation of the active site is conserved, and that it is accessible despite the GFP or 6xHis tags.

UGT76G1 produces rebaudioside C and F from steviol glycosides substrates. An in vitro activity test was carried out from a crude extract of *N. benthamiana* leaves expressing UGT76G1 and put in the presence of the Genotype E's extract obtained from dried leaves. The objective was to evaluate the activity of the enzyme in the presence of several substrates. Genotype E used to produce the SVglys extract mainly accumulates ST and Dulc A, due to the non-functionality of the β -1,3 glucosylation, as previously shown¹⁰.

Chromatograms corresponding to the negative controls show that (1) the non-transformed *N. benthamiana* leaves do not naturally produce SVglys (Fig. 6a), (2) they do not catalyze the β -1,3 glucosylation of the SVglys of the Genotype E's extract (Fig. 6c) (3) and that leaves of *N. benthamiana* expressing UGT76G1 do not produce SVglys (Fig. 6b).

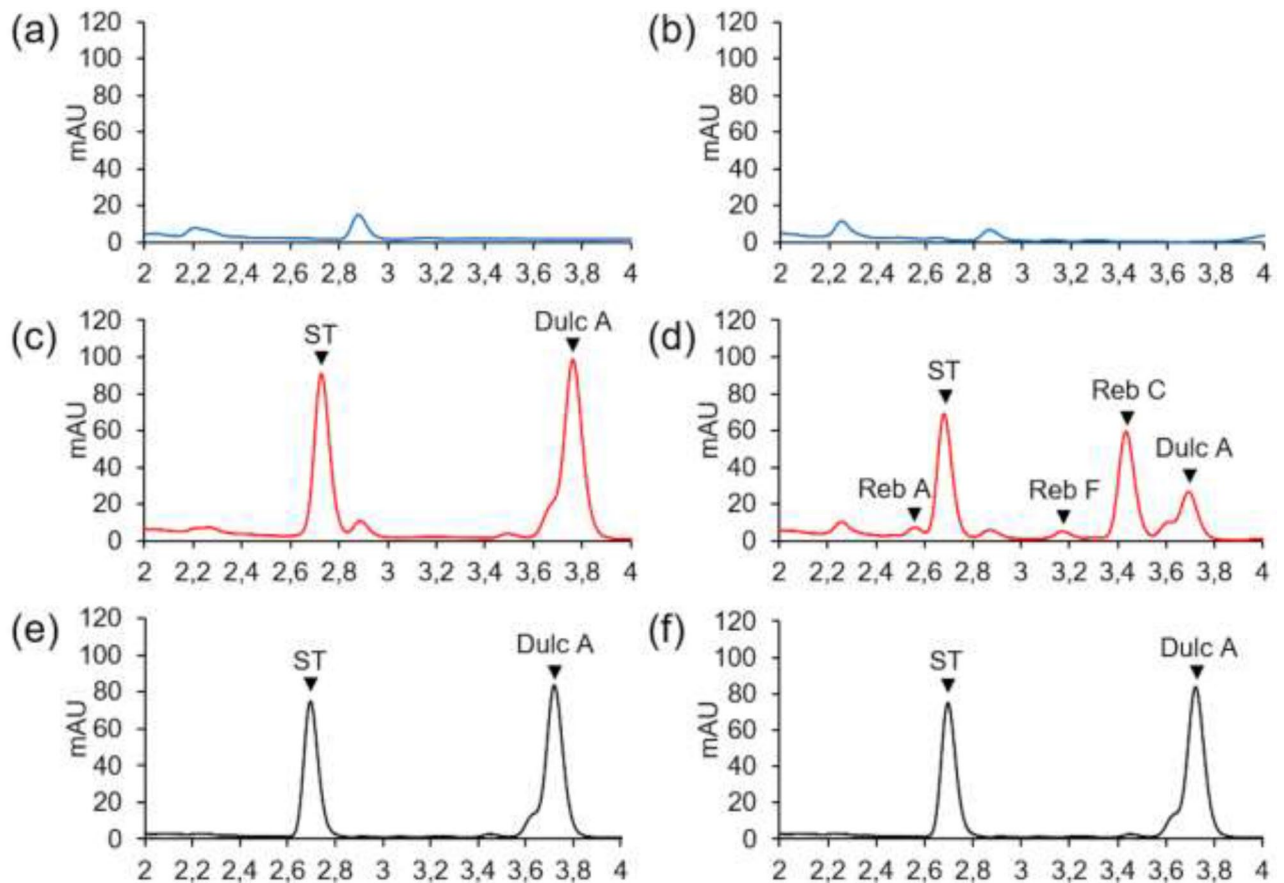


Figure 6. Functionality of UGT76G1 in presence of several SVgly substrates: in vitro activity of *N. benthamiana* crude extracts incubated without (a,b) or with (c,d) SVgly extract from Genotype E. (a,c) Non-transformed *N. benthamiana*, (b,d) *N. benthamiana* expressing UGT76G1 at 5 dpi, (e,f) SVgly extract from Genotype E. Dulc A dulcoside A, ST stevioside, Reb A rebaudioside A. Original chromatogram for (e,f) is given in Supplementary Fig. S4 and compared with SVgly standards for peak identification. Original chromatograms for (d–f) are given in Supplementary Fig. S5 for comparison of SVgly presence/absence.

Chromatograms relative to *N. benthamiana* leaves expressing UGT76G1 and put in presence of the Genotype E's extract show additional peaks with respect to ST and Dulc A (Fig. 6d). These molecules have retention times corresponding to Reb A, C and F. The surface area of the Dulc A peak decreased compared to that of the SVgly extract alone (Fig. 6e,f), in favor of an increase in the area of the Reb C peak, probably corresponding to glucosylation of Dulc A. Concerning Reb F, its precursor is not identified among SVgly peaks appearing on the chromatograms. The outgoing peak near the Dulc A seems to decrease compared to the SVgly extract alone, so it could correspond to the Reb F precursor.

Considering the proportions of Reb A and Reb C relative to those of their respective precursors, only $7.8 \pm 1.3\%$ ($n = 3$) of ST are converted into Reb A on average against $73.5 \pm 4.9\%$ of Dulc A converted in Reb C ($n = 3$), after 2 h. This isoform seems to be preferentially active with Dulc A rather than with ST. If it was the case, it could be due to easier access to the β -1,3 glycosylation site on Dulc A compared to ST. Indeed, these two compounds differ only in the nature of the β -1,2 sugar: Dulc A has a rhamnose and ST, a glucose. This hypothesis could be verified with affinity measurements on a purified protein.

Discussion

In this work, two types of activity tests were used to evaluate the functionality of a plant UGT, directly from a crude extract or plants expressing the protein of interest. These tests are based on an on/off detection system of the enzyme products and are not intended for quantification. These methods showed that recombinant UGT76G1 catalyzes the conversion of ST to Reb A, as expected¹⁴. The tests also revealed that the β -1,3 glycosylation activity of UGT76G1 was not affected by the addition of a GFP or polyhistidine tag (6xHis), both in C- and N-terminal. This allows with a single vector, to test both the activity of the enzyme and research its localization in the cell, or to purify it for a more refined enzymatic characterization.

Through expression in *N. benthamiana*, UGT76G1 was localized in cytosol, as previously shown²⁴. This compartmentalization is logical considering the role of glycosylation of UGT76G1, which occurs after synthesis of the aglycone in the chloroplast and the endoplasmic reticulum and before the storage of glycosides in the vacuole. The GFP fusions were also detected in ER which may be explained by presence of the recombinant UGT76G1

in the maturation area. Fluorescence of UGT76G1-GFP and GFP-UGT76G1 were also observed in nucleus. The same nuclear fluorescence pattern was already observed for some plant UGTs as GFP-tagged UGT85A1 from *Arabidopsis*²⁵ or GFP tagged *TaUGT3* of *Triticum aestivum* expressed in onion epidermis cells²⁶. A nuclear localization was also obtained for the stevia's UGT74G1 expressed in onion epidermal cells²⁴, for which the authors advocated for a weak diffusion of the fusion protein into the nucleus. Even if passive diffusion decreases with protein size, it was shown that proteins larger than 60 kDa passively diffuse from cytosol into the nucleus, thus accumulating over time^{27,28}. The authors found that passive diffusion does not have a molecular mass threshold, contradicting the view that the maximal size for protein diffusion through the nuclear pore is around 60 kDa^{27,28}. Another alternative to explain stevia's UGTs localization in nucleus supports the presence of a nuclear addressing and export signal although it was never shown²⁴. However, we cannot exclude the presence of the functional protein in the nucleus, as it was shown in *Zea mays* root cells for the *trans*-zeatin *O*-glycosyltransferase enzyme, for which nuclear localization was confirmed by immunodetection^{29,30}. In our case, part of the GFP fluorescence in the nucleus may also come from fusions degradation.

Many chemical phenotypes have been observed in our *S. rebaudiana* collection⁸. The diversity of SVglys profiles thus makes it possible to produce numerous extracts directly obtained from the plant, with different combinations of substrates. For example, Genotype E is not able to produce Reb A and Reb C¹⁰. Extraction of SVglys from Genotype E's leaves allow production of an extract rich in ST and Dulc A, which was used as substrates for UGT76G1. The method developed in vitro showed that UGT76G1 contained in a crude extract of *N. benthamiana* catalyzes the β -1,3 glucosylation of Dulc A to form Reb C, from Genotype E's extract. The crude extract containing UGT76G1 also produces Reb F from the *S. rebaudiana* extract, confirming the assumption previously made⁸, although the precursor of this SVgly was not identified on the chromatograms. Usually, methods to test activity in vitro require purification of the recombinant protein. Here, the developed method allows to test the functionality of several UGTs simply from a crude extract, using as substrates individual SVglys as well as SVglys synergies, moving closer to in vivo conditions. With identical crude extract/substrate ratio and reaction time, relative substrate affinity of a UGT could be investigated, by testing several individual and combined substrates. It could also be adapted to test UGTs thought to be involved in minor or higher glycosylations of SVglys, requiring a finer detection system such as mass spectrometry.

The in planta method is unique in that it infiltrates the substrate directly into the *N. benthamiana* leaf expressing the protein of interest. This method has previously been used but remains rather unusual in protein functionality studies and has never been used to study *S. rebaudiana* UGTs. For example, PgUGT5 and PgUGT5b from *Picea glauca* involved in pungenol glycosylation were expressed in *N. benthamiana*, and their substrate was directly injected into the transformed leaves³¹. However, *N. benthamiana* exhibited endogenous glycosylation activity leading to a basal production of the detected products. The use of this type of method appears to be appropriate for the study of UGTs involved in the biosynthesis of SVglys as it may limit this endogenous phenomenon. Indeed, the SVglys pathway does not exist in *N. benthamiana*, which restricts the existence of UGTs that can be recruited for the glycosylation of SVglys.

Taken together, these methods are effective tools to screen functionality of recombinant UGTs suspected to be involved in SVglys biosynthesis. Once expression vectors are produced and *N. benthamiana* is transformed, the user can choose the best method. The in vitro method only requires one expression vector and is suited to quickly screen several substrates and non-characterized UGTs. The in planta method uses binary vectors from pGWBs series and allows simultaneously functional screening close to the in vivo conditions, and several applications for further analysis, such as subcellular localization or purification.

Material and methods

Biological material. *Nicotiana benthamiana*. *Nicotiana benthamiana* seeds used for transient transformation were provided by LRSV (Toulouse, France). Each seed was sown in a 7 × 7 cm plastic pot in a special seedling and cuttings potting soil. The seeds were sprouted in moisture-saturated plastic mini-greenhouses (closed aeration) to allow germination and good plant development. The seedlings were placed under long day conditions: 16 h of day and 8 h of night, at a temperature of 20–23 °C. After the appearance of the first 3 leaves, the vents of the mini-greenhouses were opened (weekly watering) then the lids of the mini-greenhouses removed (daily watering), in order to accustom the plants to the ambient hygrometry.

Stevia rebaudiana. Our germplasm is derived from selection cycles of Criolla and Morita populations. Genotype E previously studied, has a non-functional UGT76G1¹⁰ and doesn't produce Reb A and Reb C, and accumulates high proportions of Dulc A and ST. Genotype E was cultivated in I-L pot in greenhouse under long days (16 h of daylight), to maintain plants in a vegetative stage and favor biomass production.

Production of expression vectors. The coding sequence of genotype A's UGT76G1 of our previous study was used for the experiments¹⁰. Primers used for coding sequence amplification of UGT76G1 were used to design attB specific primers, for native expression and C- and N-term fusions (Table 1).

The synthesis of attB-UGT76G1 PCR products was performed in two steps. Platinum Taq DNA Polymerase High Fidelity (Thermo Fisher Scientific) was used for the two PCR, according to manufacturer recommendations. The first PCR (50 μ L) consisted in amplifying purified clones containing UGT76G1 of the Genotype A with the specific primers of Table 1. The cycling parameters described in Table 2 were used.

Ten microliters of the first PCR reaction were transferred to a 40 μ L PCR mixture containing 40 pmol each of the attB1 and attB2 adapter primers (AttB1: GGGGACAAGTTTGTACAAAAAAGCAGGCT, AttB2: GGGG ACC ACTTTGTACAAGAAAGCTGGGT.). The second PCR (50 μ L) used the cycling parameters described in Table 3.

Orientation	Primer	Usage
Forward	AAAAAGCAGGCTATGGAAAATAAACGGAGACCA	Native expression
	AAAAAGCAGGCTTCATGGAAAATAAACGGAGACCA	C-term fusion
Reverse	AGAAAGCTGGGTTTACAACGATGAAATGTAAGAACTAG	Native expression
	AGAAAGCTGGGTTTACAACGATGAAATGTAAGAACTAG	N-term fusion
	AGAAAGCTGGGTGCAACGATGAAATGTAAGAACTAG	C-term fusion

Table 1. UGT76G1 specific primers used to synthesize attB-PCR products.

Denaturation	2 min	95 °C	
Denaturation	15 s	94 °C	10 cycles
Annealing	30 s	53.3 °C	
Elongation	1 min/kb	68 °C	

Table 2. Cycling parameters for the first step PCR.

Denaturation	2 min	95 °C	
Denaturation	15 s	94 °C	5 cycles
Annealing	30 s	45 °C	
Elongation	1 min/kb	68 °C	
Denaturation	15 s	94 °C	15–20 cycles
Annealing	30 s	55 °C	
Elongation	1 min/kb	68 °C	
Elongation	10 min	68 °C	

Table 3. Cycling parameters for the second step PCR.

Entry clones were created by BP recombination between the attB-UGT76G1 purified products and the vector pDONR/Zeo (Invitrogen, Thermo Fisher Scientific) with the BP Clonase enzyme (Invitrogen, Thermo Fisher Scientific), following the manufacturer recommendations. Library Efficiency DH5 α (Thermo Fisher Scientific) competent cells were transformed by heat shock at 42 °C during 45 s with the entry clones. Positive clones were selected and sequenced to confirm the insertion of the transgene and the correct reading frame for fusion proteins. Entry clones were then purified with the Wizard Plus Minipreps DNA Purification System (Promega) according to the manufacturer recommendations. Expression clones were produced by a LR recombination between the entry clones and the binary vectors pGWBs³², with the LR Clonase enzyme (Invitrogen, Thermo Fisher Scientific). The vectors pGWBs were provided by T. Nakagawa (Shimane University, Japan). Library Efficiency DH5 α (Thermo Fisher Scientific) competent cells were transformed by heat shock at 42 °C during 45 s with the expression clones. Positive clones were selected and purified with the purification system NucleoBond Xtra Midi (Macherey–Nagel) following the manufacturer recommendation.

***Agrobacterium tumefaciens* transformation.** The LBA4404 strain was used for transformation (ElectroMAX *A. tumefaciens* LBA4404 Cells, Invitrogen, Thermo Fisher Scientific). Cells were transformed with expression vectors in electroporation cuvettes, during 5 ms at 1.8 kV, 200 Ω , 25 μ F. After electroporation, 800 μ L of YEB medium were added and cultures were incubated at 28 °C during 2 h at 180 rpm. Cells were spread on LB plates with rifampicin, streptomycin, and hygromycin (at 100 μ g/mL final) and incubated during 48 h at 28 °C. Positive clones were selected and pricked out in LB liquid medium.

***Nicotiana benthamiana* transformation.** For each construct, a 200 mL-culture with appropriate antibiotics was prepared starting from 40 μ L of a glycerol stock. Cells were grown at 28 °C, 150 rpm until $1.8 < DO_{600nm} < 2.0$. They were then centrifuged 20 min at 1,000g. Media was removed and cells were washed with the resuspension solution (10 mM MgCl₂, 10 mM MES-K pH 5.6). After wash, pellets were resuspended in the resuspension solution containing acetosyringone 100 μ M. An aliquot of the bacterial solution was sampled and diluted in the resuspension solution with acetosyringone qsp. the desired volume until $0.7 < DO_{600nm} < 0.9$. Cells were incubated during 2 h at ambient temperature. *Agrobacterium* was infiltrated into *N. benthamiana* leaves with a 5 mL-syringe applied on the lower face. As control of plant transformation, *A. tumefaciens* LBA4404 strain containing a vector pEAQ-HT-Dest1-GFP was used³³. This strain was provided by L. Hoffman (University

of Toulouse, France). Vectors containing *UGT76G1* and *GFP* were co-infiltrated with the co-suppressor p19 (with a bacterial solution volume ratio 2:1) as it enhances recombinant protein expression (see Supplementary Fig. S6). The vector pCB301-p19 was provided in a LBA4404 strain by J. Win (Sainsbury Laboratory, United-Kingdom).

Protein analysis. *Sample preparation.* Leaves were sampled 5 days post infiltration. Immediately after sampling, leaves were frozen in liquid nitrogen. They were ground into a fine powder with a ball mill MM300 (Retsch) and stainless-steel balls (2 mm Ø). One volume of sample was extracted in 2 volumes of extraction buffer (Triton X100 0.25%, DTT 10 mM, Halt Protease Inhibitor Cocktail [Thermo Fisher Scientific] 1X, qsp. GTEN buffer (glycerol 10%, Tris-HCl pH 7.5 150 mM, EDTA 1 mM, NaCl 150 mM)]. Samples were incubated 10 min at 4 °C under gentle stirring. After centrifugation 10 min at 4 °C, 3,000g, supernatants were sampled and conserved for protein analysis. Proteins were quantified using the Bradford method. For western blot, 20 µg of proteins were denatured 5 min at 96 °C using 5 volumes of protein extract and 1 volume of Laemmli 6X. Samples were centrifugated 30 s at 13,000g and supernatants were used for electrophoresis.

Western blot. Proteins were loaded on a 1 mm gel prepared with TGX Stain-Free FastCast Acrylamide Kit, 10% (BioRad) according to manufacturer specifications. Proteins were separated in a Mini-Protean Tetra Cell System (BioRad) at 250 V (400 mA) during 30 min. Separated proteins were transferred on nitrocellulose membrane with a Trans-Blot Turbo Transfer System (BioRad) at 25 V during 3 min per gel. Membranes were saturated during 1 h under strong stirring with 5% fat free milk dissolved in TBS Tween 0.1%. Then, they were incubated with primary antibodies prepared at 1/1,000 in BSA 1X TBS-Tween 1X during 2 h at ambient temperature. For blotting of GFP-tagged proteins, monoclonal anti-GFP antibodies produced in mouse were used respectively (Sigma). After incubation with primary antibodies, membranes were washed three times with TBS-Tween 1X during 10 min and incubated under stirring during 1 h with secondary antibodies prepared at 1/6,000 in BSA 1X TBS Tween 1X. Secondary antibodies anti-mouse IgG (H + L)-HRP conjugated antibodies produced in goat were used (BioRad). Membranes were washed three times with TBS-Tween 1X during 10 min. Specific bands were revealed with Clarity Western ECL Substrate (BioRad) and membranes were observed with a ChemiDoc XRS+ system (BioRad). For the revelation step, anti-GFP blots were exposed for 60s except for GFP controls (3s) to avoid overexposition. The 60s exposition for the GFP control is shown on Supp fig 1.

Microscopy. Microscopy observations were realized at the platform FR AIB (Toulouse, France) at 5 days post-infiltration. GFP fluorescence ($\lambda_{\text{excitation}} = 488 \text{ nm}$; $\lambda_{\text{emission}} = 510 \text{ nm}$) was observed with a confocal laser scanning microscope (CLSM) TCS SP8 (Leica) and a HC Fluotar L25X/0.95 W VISIR (Leica) lens. Pictures were taken with the software LAS (Leica) and were processed with the software ImageJ (1.52a).

Activity of recombinant proteins. *Production of the SVglys extract.* Genotype E's leaves (complete stems) were sampled and dried at 50 °C. Then, they were ground with a mortar and pestle. SVglys were extracted from 10 g of Genotype E's dry leaves during 30 min at 50 °C under stirring in 200 mL of ultrapure water. The extract was filtrated with absorbent paper to remove leaf powder and fragments.

In planta: infiltration of substrate in transformed leaves. At 5 dpi, transformed leaves were infiltrated by syringe with stevioside 5 mM. After 24 h, leaves were sampled separately and dried 48 h at 50 °C. Then, they were grinded with piston pellet Eppendorf and steviol glycosides were extracted 30 min at 50 °C in 200 µL of ultrapure water. Extracts were filtrated at 0.2 µM and analyzed by UPLC.

In vitro: enzymatic assay on crude extract expressing recombinant protein. For each construct, two transformed leaves were sampled at 5 dpi and immediately ground in ice with 500 µL of cold reaction buffer (Tris-HCl 100 mM, pH 8.0; MgCl₂ 3 mM; BSA 10 µg mL⁻¹). Extracts were centrifugated 2 min at 2,000g at 4 °C, and 180 µL of supernatant were transferred in two new tubes. Twenty microliters of reaction buffer were added in one of the tubes to constitute the negative control. Twenty microliters of substrate were added in the other tube. The substrates were alternatively stevioside 5 mM or SVglys extract for the Genotype E, mainly composed of ST and Dulc A.

Tubes were homogenized by gentle flipping. After 2 h at 30 °C, reactions were stopped by addition of 10 µL TFA 20%. Samples were centrifugated 5 min at 2,000g. Supernatants were filtrated at 0.2 µm and analyzed by UPLC.

UPLC analysis of steviol glycosides. Liquid chromatography was performed under isocratic conditions on a Dionex UltiMate 3,000 system from ThermoScientific consisting of a DGP-3600RS pump, a WPS-3000 TRS autosampler, a TCC-3000 SD column compartment maintained at 40 °C and a DAD-3000 diode array detector set to a wavelength of 195 nm. The flow rate and sample injection volume were 0.3 mL min⁻¹ and 1 µL, respectively. Steviol glycosides were separated with a reverse phase (RP) column Kinetex C18 (150 × 2.1 mm, 1.7 µm) (Phenomenex, USA), maintained at 40 °C. Solvents used as mobile phase were 32% acetonitrile and 68% ultra-high-purity water acidified with 0.1% formic acid (pH 2.6). A mix of 9 SVglys standards (JECFA mixture) was purchased from Chromadex, Inc. (Irvine, CA, USA) and were used to identify ST, Dulc A, Reb A and C. Chromatogram of standards and UV spectrum of ST are given in the Supplementary Fig. S2.

Received: 28 January 2020; Accepted: 12 August 2020

Published online: 15 September 2020

References

1. Ceunen, S. & Geuns, J. M. C. Steviol glycosides: chemical diversity, metabolism, and function. *J. Nat. Prod.* **76**, 1201–1228 (2013).
2. Cheng, A.-X. *et al.* Plant terpenoids: biosynthesis and ecological functions. *J. Integr. Plant Biol.* **49**, 179–186 (2007).
3. Totté, N. *et al.* Biosynthesis of the diterpenoid steviol, an ent-kaurene derivative from *Stevia rebaudiana* Bertoni, via the methylerythritol phosphate pathway. *Tetrahedron Lett.* **41**, 6407–6410 (2000).
4. Totté, N. *et al.* Corrigendum to “Biosynthesis of the diterpenoid steviol, an ent-kaurene derivative from *Stevia rebaudiana* Bertoni, via the methylerythritol phosphate pathway”: [Tetrahedron Letters 41 (2000) 6407]. *Tetrahedron Lett.* **41**, 7595 (2000).
5. Brandle, J. E. & Telmer, P. G. Steviol glycoside biosynthesis. *Phytochemistry* **68**, 1855–1863 (2007).
6. Gachon, C. M. M., Langlois-Meurinne, M. & Saindrenan, P. Plant secondary metabolism glycosyltransferases: the emerging functional analysis. *Trends Plant Sci.* **10**, 542–549 (2005).
7. Abdelsalam, N. R. *et al.* Comparison of uridine diphosphate-glycosyltransferase *UGT76G1* genes from some varieties of *Stevia rebaudiana* Bertoni. *Sci. Rep.* **9**, 1–10 (2019).
8. Barbet-Massin, C., Giuliano, S., Alletto, L., Daydé, J. & Berger, M. Towards a semi-perennial culture of *Stevia rebaudiana* (Bertoni) Bertoni under temperate climate: effects of genotype, environment and plant age on steviol glycoside content and composition. *Genet. Resour. Crop Evol.* **63**, 685–694 (2016).
9. Francisco, F., Pereira, G. P., Machado, M. P., Kanis, L. A. & Deschamps, C. Characterization of *Stevia rebaudiana* Bertoni accessions cultivated in Southern Brazil. *J. Agric. Sci.* **10**, 353 (2018).
10. Petit, E., Jacques, A., Daydé, J., Vallejo, V. & Berger, M. *UGT76G1* polymorphism in *Stevia rebaudiana*: New variants for steviol glycosides conjugation. *Plant Physiol. Biochem.* **135**, 563–569 (2019).
11. Yang, Y. *et al.* Base substitution mutations in uridinediphosphate-dependent glycosyltransferase *76G1* gene of *Stevia rebaudiana* causes the low levels of rebaudioside A: Mutations in *UGT76G1*, A key gene of steviol glycosides synthesis. *Plant Physiol. Biochem.* **80**, 220–225 (2014).
12. Richman, A. S., Gijzen, M., Starratt, A. N., Yang, Z. & Brandle, J. E. Diterpene synthesis in *Stevia rebaudiana*: recruitment and up-regulation of key enzymes from the gibberellin biosynthetic pathway. *Plant J. Cell Mol. Biol.* **19**, 411–421 (1999).
13. Zheng, J., Zhuang, Y., Mao, H.-Z. & Jang, I.-C. Overexpression of *SrDXS1* and *SrKAH* enhances steviol glycosides content in transgenic *Stevia* plants. *BMC Plant Biol.* **19**, 1 (2019).
14. Richman, A. *et al.* Functional genomics uncovers three glucosyltransferases involved in the synthesis of the major sweet glucosides of *Stevia rebaudiana*. *Plant J.* **41**, 56–67 (2005).
15. Li, W. *et al.* Development of photoaffinity probe for the discovery of steviol glycosides biosynthesis pathway in *Stevia rebaudiana* and rapid substrate screening. *ACS Chem. Biol.* **13**, 1944–1949 (2018).
16. Vogt, T. & Jones, P. Glycosyltransferases in plant natural product synthesis: characterization of a supergene family. *Trends Plant Sci.* **5**, 380–386 (2000).
17. Osmani, S. A., Bak, S. & Møller, B. L. Substrate specificity of plant UDP-dependent glycosyltransferases predicted from crystal structures and homology modeling. *Phytochemistry* **70**, 325–347 (2009).
18. Fukuchi-Mizutani, M. *et al.* Biochemical and molecular characterization of a novel UDP-glucose:anthocyanin 3'-O-glucosyltransferase, a key enzyme for blue anthocyanin biosynthesis, from gentian. *Plant Physiol.* **132**, 1652–1663 (2003).
19. Dewitte, G. *et al.* Screening of recombinant glycosyltransferases reveals the broad acceptor specificity of stevia *UGT-76G1*. *J. Biotechnol.* **233**, 49–55 (2016).
20. Terpe, K. Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems. *Appl. Microbiol. Biotechnol.* **60**, 523–533 (2003).
21. Weill, U. *et al.* Assessment of GFP tag position on protein localization and growth fitness in yeast. *J. Mol. Biol.* **431**, 636–641 (2019).
22. Seibel, N. M., Eljouni, J., Nalaskowski, M. M. & Hampe, W. Nuclear localization of enhanced green fluorescent protein homomultimers. *Anal. Biochem.* **368**, 95–99 (2007).
23. Wei, X., Henke, V. G., Strübing, C., Brown, E. B. & Clapham, D. E. Real-time imaging of nuclear permeation by EGFP in single intact cells. *Biophys. J.* **84**, 1317–1327 (2003).
24. Humphrey, T. V., Richman, A. S., Menassa, R. & Brandle, J. E. Spatial organisation of four enzymes from *Stevia rebaudiana* that are involved in steviol glycoside synthesis. *Plant Mol. Biol.* **61**, 47–62 (2006).
25. Jin, S.-H. *et al.* Overexpression of glucosyltransferase *UGT85A1* influences trans-zeatin homeostasis and trans-zeatin responses likely through O-glucosylation. *Planta* **237**, 991–999 (2013).
26. Lulin, M. *et al.* Molecular cloning and characterization of an up-regulated UDP-glucosyltransferase gene induced by DON from *Triticum aestivum* L. cv. Wangshuibai. *Mol. Biol. Rep.* **37**, 785 (2009).
27. Timney, B. L. *et al.* Simple rules for passive diffusion through the nuclear pore complex. *J. Cell Biol.* **215**, 57–76 (2016).
28. Wang, R. & Brattain, M. G. The maximal size of protein to diffuse through the nuclear pore is larger than 60kDa. *FEBS Lett.* **581**, 3164–3170 (2007).
29. Li, R., Sosa, J. L. & Zavala, M. E. Accumulation of zeatin O-glycosyltransferase in *Phaseolus vulgaris* and *Zea mays* following cold stress. *Plant Growth Regul.* **32**, 295–305 (2000).
30. Šmečilová, M., Dobrušková, J., Novák, O., Takáč, T. & Galuszka, P. Cytokinin-specific glycosyltransferases possess different roles in cytokinin homeostasis maintenance. *Front. Plant Sci.* **7**, 1264 (2016).
31. Mageroy, M. H. *et al.* A conifer UDP-sugar dependent glycosyltransferase contributes to acetophenone metabolism and defense against insects. *Plant Physiol.* **175**, 641–651 (2017).
32. Nakagawa, T. *et al.* Development of series of gateway binary vectors, pGWBs, for realizing efficient construction of fusion genes for plant transformation. *J. Biosci. Bioeng.* **104**, 34–41 (2007).
33. Sainsbury, F., Thuenemann, E. C. & Lomonosoff, G. P. pEAQ: versatile expression vectors for easy and quick transient expression of heterologous proteins in plants. *Plant Biotechnol. J.* **7**, 682–693 (2009).

Acknowledgements

This work was funded by PepsiCo Global R&D. The views expressed in this article are those of the authors and do not necessarily reflect the position or policy of PepsiCo Inc. The authors are grateful to Lauriane Heissat for production and culture of *Stevia rebaudiana* and to Apolline Champenois for her help in experiments.

Author contributions

M.B., A.J., and E.P. designed the study. E.P. carried out the experiments, analyzed the data, and wrote the manuscript. L.C. provided support on transformation and western blot. M.B., L.C., J.D., A.J., E.P., and V.V. interpreted the results and contributed to the manuscript.

Competing interests

V. Vallejo is employed by PepsiCo, Inc. A. Jacques, M. Berger, L. Camborde, J. Daydé, E. Petit and V. Vallejo declare no potential conflict of interest.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-71746-9>.

Correspondence and requests for materials should be addressed to M.B.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020