CENTRO DE INVESTIGACIÓN Y DE ESTUDIOS AVANZADOS
DEL INSTITUTO POLITÉCNICO NACIONAL
UNIDAD IRAPUATO

Análisis de líneas transgénicas de *Arabidopsis thaliana* expresando el cDNA sacarosa: sacarosa 1-fructosiltransferasa (*Atq1SST-1*) de *Agave tequilana*.

Tesis que presenta

IA. Alan Daniel Gomez Vargas

Para obtener el grado de

Maestro en ciencias

En la especialidad de

Biotecnología de plantas

Director de tesis

Dra. June Kilpatrick Simpson Williamson

Irapuato, Guanajuato **Noviembre, 2018**

CENTRO DE INVESTIGACIÓN Y DE ESTUDIOS AVANZADOS
DEL INSTITUTO POLITÉCNICO NACIONAL
UNIDAD IRAPUATO

Analysis of transgenic *Arabidopsis thaliana* lines expressing a sucrose: sucrose 1-fructosyltransferase (*Atq1SST-1*) cDNA from *Agave tequilana*.

Thesis presented by

Engineer Alan Daniel Gomez Vargas

To obtain the grade of

Master of Science in Plant Biotechnology

Thesis advisor

Ph.D. June Kilpatrick Simpson Williamson

Irapuato, Guanajuato **November 2018**

The current research was developed at the department of Genetic Engineering, Laboratory of Molecular Genetics of Sexual and Asexual Development (CINVESTAV Irapuato), under the direction of Ph.D. June Kilpatrick Simpson Williamson.

Acknowledgments

First of all, I would like to make a special acknowledgment to the person that four years ago let me be part of this incredible work team and made me become the person that I am today, Dr. June Simpson thank you for being such a good mentor and for always supporting me and believing in me even though most of the time I thought that I could not make it. Thank you for sharing your time, knowledge, advice and for insisting to keep trying in my failed experiments. At the end, I persevered from my failures and I realized that I had done things that I never thought that I would be capable of. Thanks to you I finally found a scientific inspiration in my life.

Thanks to M.Sc. (almost PhD) Emmanuel Avila de Dios for being such an incredible teacher and mentor, first for my undergraduate thesis, and now once again for my master thesis. I will always be deeply grateful for all the time and effort you put on me by teaching me the theoretical and practical explanations of molecular biology throughout all these years. I have to be honest, all the things that I learned in all these years was because of you. Lastly, I will not forget something that you said to me in the first days of my arrival in the lab and it is that the most important part of an experiment is not the fact of following a protocol, or a recipe like you commonly say, but knowing the scientific basis of the method. These words I will keep in my mind forever.

To M. Sc (almost PhD) Rocio Aguilar for all the help provided in some experiments, and especially during the quantitative PCR analysis. Thank you for solving all my doubts at a distance. Now I can proudly say that I am capable of performing a qPCR as perfectly, as you used to perform them. Obviously, I learnt from the best in the area. Thank you for your real friendship throughout all these years, it means a lot to me having you as a friend. Finally, thank you for all the priceless moments we shared together.

To PhD. John Paul Délano Frier and PhD. Stefan de Folter for being part of this project, as well as the advices and contributions given throughout the experimental phase. Also, I would like to thank PhD. Plinio Guzmán for taking the place of PhD. Stefan at the last minute.

To the “Consejo Nacional de Ciencia y Tecnología” (CONACyT) for the scholarship provided for the master program 608587 and financial support under the project Ciencia Básica 2013-220339.

To Cinvestav Unidad Irapuato and all its administrative staff.

To PhD. Jazmin Abraham Juarez for all the help provided at a distance with my qPCR experiments. I hope one day I can become as good as you are at science.

To M. Sc Katia del Carmen Gil Vega for all the technical support and assistance throughout the experimental phase, also for the advice and comments received and for being so efficient in getting me all the material required during the masters and for being such a good friend.

To M. Sc. Arely Pérez for all the advices and suggestions during my experiments, seminars, etc., for all the explanations you taught me about Maldi analysis and how to interpret the result (well I still need to improve that) and for carrying out Maldi analysis in some of my samples.

To Mom, Dad and my siblings for being such an incredible family. Thank you for all the support in my most frustrating time during the master and for not letting me quit, you always had faith in me even when I thought I was a complete failure. Mom and Dad, I am so proud of you, you did not have the opportunity of having any type of academic degree, but now, thanks to you I can become whoever I want to be and the same for my siblings.

To Lina Lopez for being one of my greatest friends. I had such an amazing time in all the thing that we have been through during our trips. It was a pleasure having you as a lab mate.

To Erika Bautista for letting me know your personal life, at the end we have so much in common, and I think that is the reason that we have a strong friendship.

To Laura Zavala for 7 years of friendship and for always having you in the most important events of my life and also for letting me be part of yours as well. I will always appreciate your confidence that you always put on me.

To Simpson´s lab members: Katia, Laura, Emmanuel, Chio, Arely, Erika, Lina, Gerardo (aka Chino), Yoselin and Andrea, you guys are amazing.

Agradecimientos

Primero que nada, me gustaría hacer un agradecimiento especial a la persona que cuatro años atrás me permitió ser parte de este increíble equipo de trabajo y me hizo ser la persona que hoy soy, Dra. June Simpson gracias por ser una gran mentora y por apoyarme y creer en mi, aunque la mayor parte del tiempo pensé que no podría lograrlo. Gracias por compartir su tiempo, conocimiento, consejos y por insistirme en seguir intentando mis experimentos fallidos, al final, perseveré de mis fracasos y pude darme cuenta de que había hecho cosas que nunca pensé que sería capaz de hacerlas. Gracias a usted finalmente pude encontrar una inspiración científica en mi vida.

Gracias al maestro en ciencias (casi doctor) Emmanuel Ávila de Dios por ser un increíble maestro y mentor primero en la tesis de licenciatura y una vez mas en la tesis de maestría. Siempre estaré profundamente agradecido por todo el tiempo y esfuerzo que me has brindado al enseñarme todas las explicaciones teóricas y prácticas de la biología molecular a lo largo de todos estos años. Tengo que ser honesto, todas las que cosas que aprendí en todos estos años fue gracias a ti. Finalmente, nunca olvidaré algo que me dijiste en los primeros días que llegué al laboratorio y es que la parte mas importante de un experimento no es el hecho de seguir un protocolo, o una receta como comúnmente dices, sino saber las bases científicas del método. Esas palabras permanecerán en mi mente por siempre.

A la maestra en ciencias (casi doctora) Rocío Aguilar por toda la ayuda brindada en algunos de mis experimentos, y especialmente durante los análisis de PCR cuantitativo. Gracias por resolver todas mis dudas a la distancia, ahora puedo decir orgullosamente que tengo la capacidad de realizar qPCR casi tan perfectos como tu solías hacerlos. Obviamente, aprendí de la mejor en el área. Gracias por tu amistar verdadera a lo largo de estos años, significa mucho para mi tenerte como amiga. Finalmente, gracias por todos los momentos invaluables que compartimos.

A los Doctores John Paul Délano Frier y Stefan de Folter por querer ser parte de este proyecto, así como todos los consejos y aportaciones dadas durante la parte experimental. Así mismo, me gustaría agradecer al Dr. Plinio Guzmán por ocupar el lugar del Dr. Stefan a último momento.

Al Consejo Nacional de Ciencia y Tecnología (CONACyT) por la beca otorgada para el programa de maestría 608587 y por el soporte financiero bajo el proyecto de Ciencia Básica 220339.

A Cinvestav unidad Irapuato y todo el personal administrativo.

A la Dra. Jazmín Abraham Juárez por toda la ayuda otorgada a la distancia con mis experimentos de qPCR. Espero un día pueda ser tan bueno como tú en la ciencia.

A la maestra en ciencias Katia del Carmen Gil Vega por todo el soporte técnico y asistencia durante la fase experimental, así mismo por los consejos y comentarios recibidos y por ser tan eficiente en obtenerme todo el material que requería durante la maestría y por ser una gran amiga.

A la maestra en ciencias Arely Pérez por todos los consejos y sugerencias durante mis experimentos, seminarios, etc., por todas las explicaciones que me enseñaste acerca de los análisis de Maldi y como interpretar los resultados (aunque aún necesito mejorar esa parte) y por realizar los análisis de Maldi en algunas de mis muestras.

A mi mamá, papá y hermanos por ser siempre una familia muy increíble. Gracias por todo el apoyo en mis momentos de mayor frustración durante la maestría y por no dejarme renunciar, siempre tuvieron fe en mi incluso cuando llegué a pensar que era un completo fracaso. Mamá y papá, estoy muy orgulloso de ustedes, ustedes no tuvieron la oportunidad de adquirir algún tipo de grado académico, pero ahora, gracias a ustedes puedo ser quien sea que quiera ser, siendo lo mismo para mis hermanos.

A Lina López por ser una de mis mejores amigas. Pasé momentos increíbles en todo lo que hemos pasado durante los viajes. Fue un placer tenerte como compañera de laboratorio.

A Erika Bautista por dejarme conocer tu vida personal, al final compartimos muchas cosas en común, y creo que es esa la razón por la cual tenemos una amistad tan sólida.

A Laura Zavala por 7 años de amistad y por siempre tenerte en los momentos mas importantes de mi vida y por dejarme ser parte de los tuyos también. Siempre apreciaré la confianza que me das.

A los miembros del Simpson´s lab: Katia, Laura, Emmanuel, Chio, Arely, Erika, Lina, Gerardo (también conocido como “chino”), Yoselin y Andrea, todos ustedes son increíbles.

Abstract

Fructans are fructose polymers synthesized from sucrose molecules and by the catalytic activity of enzymes known as fructosyltransferases (FT). These polysaccharides are the main carbohydrate storage molecules in 15% of angiosperms, including *Agave tequilana*. Fructans synthesized by *A. tequilana,* known as agavins, have been characterized at the structural and chemical level, being described as highly complex molecules. At the molecular level, several FT genes have been identified and some isoforms which theoretically should be expressed and active in order to synthesize complex fructan structures such as the agavins, have been characterized at the functional level in *A. tequilana.*

Additionally, there are reports describing the presence of fructan exohydrolases (FEH), enzymes, whose catalytic activity is characterized by fructan molecule breakdown in plants that naturally do not synthesize and accumulate fructans such as *Arabidopsis thaliana* (Arabidopsis). Through cloning and subsequent heterologous expression in the *Pichia pastoris* system, it was possible to purify, confirm and characterize at the functional level two enzymes with FEH activity, that initially were annotated as putative cell wall invertases. However, the biological function that these enzymes could be performing is not yet known, because the FEH substrate is not synthesized endogenously in Arabidopsis. Finally, there is no experimental gene expression data available from tissues or development stages of the plant that could suggest a role for FEHs in Arabidopsis.

In this work, Arabidopsis Col. 0 transgenic lines expressing the cDNA encoding the FT sucrose: sucrose 1-fructosyltransferase of *A. tequilana* (*Atq1SST-1*) were produced. The transgenic lines were analyzed to confirm *Atq1SST-1* expression and presence of fructans by thin layer chromatography (TLC). The effect of ectopic expression of *Atq1SST-1* on the endogenous *FEH* genes expression levels was analyzed and was shown to have a negative effect on *FEH* transcripts accumulation.

Resumen

Los fructanos son polímeros de fructosa sintetizados a partir de moléculas de sacarosa y por la actividad catalítica de enzimas conocidas como fructosiltransferasas (FT). Estos polisacáridos son los principales carbohidratos de reserva en 15% de plantas angiospermas, incluyendo *Agave tequilana.* Los fructanos sintetizados por *A. tequilana*, conocidos como agavinas, han sido caracterizados a nivel estructural y químico, siendo descritos como moléculas altamente complejas. A nivel molecular, varios genes FT han sido identificados y algunas isoformas que teóricamente deberían de estar expresadas y activas con el fin de sintetizar las estructuras tan complejas de los fructanos como lo son las agavinas, han sido caracterizados a nivel funcional en *A. tequilana*.

Adicionalmente, se tienen reportes describiendo la presencia de fructo exohidrolasas (FEH), enzimas cuya actividad catalítica es caracterizada por la hidrólisis de moléculas de fructanos, en plantas que naturalmente no sintetizan y acumulan fructanos tal como lo es *Arabidopsis thaliana* (Arabidopsis). Mediante clonación y subsecuente expresión heteróloga en *Pichia pastoris*, fue posible purificar, confirmar y caracterizar a nivel funcional dos enzimas con actividad FEH, que inicialmente fueron anotadas como invertasas de pared celular putativas. Sin embargo, la función biológica que pudieran desempeñar estas enzimas es desconocida todavía, porque el sustrato de las FEH no es sintetizado endógenamente en Arabidopsis. Finalmente, aun no se tienen datos de expresión génica disponibles de tejidos y etapas de desarrollo de la planta que pudieran sugerir el posible papel de las FEH en Arabidopsis.

En el presente trabajo, líneas transgénicas de Arabidopsis Col. 0 expresando un cDNA que codifica la FT sacarosa: sacarosa 1-fructosiltransferasa de *A. tequilana* (*Atq1SST-1*) fueron generadas. Las líneas transgénicas fueron analizadas para confirmar la expresión de *Atq1SST-1* y determinar la presencia de fructanos mediante cromatografía en capa fina (TLC, por sus siglas en inglés). El efecto de la expresión ectópica del transgén *Atq1SST-1*

sobre los niveles de expresión de los genes endógenos de *FEH* fue analizado, mostrando un efecto negativo sobre la acumulación de los transcritos de las FEH.

Contents

List of figures

List of tables

Introduction

The most common non-structural carbohydrates present in biological systems of the Plantae kingdom are polymers based on glucose molecules such as starch. However, another type of non-structural carbohydrates composed of fructose-based oligo and polysaccharides are present in approximately 15% of angiosperm plants (Hendry, 1993), and are known as fructans. Fructans are considered as storage carbohydrates that are used as carbon and energy sources to perform steady growth and development of plants, acting as a long-term storage carbohydrate in certain organs or as a short-term reserve carbohydrate, depending on the species (Versluys, *et al*., 2018). Although, several other functions have been described and related to fructan molecules, including membrane stabilization and free radical scavenging throughout abiotic stress conditions (Livingston, *et al*., 2009; Valluru and Van den Ende, 2008).

As mentioned above, fructans are fructose-based oligo and polysaccharides that are synthesized within the vacuolar compartment, where transference of fructose moieties from sucrose to sucrose or fructan molecules takes place, by the catalytic activity of fructosyltransferase (FT) enzymes. In monocotyledons, there are four fructosyltransferases described that are involved in fructan biosynthesis: sucrose: sucrose 1-fructosyltransferase (1-SST), sucrose: fructan 6-fructosyltransferase (6-SFT), fructan: fructan 6G-fructosyltransferase (6G-FFT) and fructan: fructan 1-fructosyltransferase (1-FFT). Fructans are categorized in several groups according to the glycosidic linkage formed to join the fructose moieties, which can be β-(2, 1) or β-(2, 6), and the molecule structure can be either linear, branched or both. Lastly, fructan breakdown is carried out by fructan exohydrolase (FEH) enzymes. These enzymes, along with vacuolar invertases and cell wall invertases, are members of Plant Glycoside Hydrolase Family 32 (PGHF32) (Vijn and Smeekens, 1999; Banguela and Hernandez, 2006; Yildiz, 2010; Cimini, *et al.,* 2015).

Agave tequilana is a monocotyledon, perennial and monocarpic plant included within the 15% of angiosperm plants capable of synthesizing fructans as the sole storage carbohydrate. The agave stem (known as "piña") is the organ where the highest concentration of fructans is stored throughout the agave life cycle (the life span of *A. tequilana* is between 6 to 8 years). Thus, *A. tequilana* is of an enormous economic value across the Mexican territory due to the use of the stems as the only raw material permitted for the production of tequila, an alcoholic beverage. In addition to *A. tequilana*, other agave species are used as a source of food or beverages (e. g., alcoholic beverages such as mescal or sotol, just to mention a few, and "aguamiel", a natural juice rich in polysaccharides generated in some agaves stems), and also as tools, natural fibers, material for construction, etc. since both ancient and modern Mexicans have exploited them for the same purposes (Nava-Cruz, *et al.,* 2015).

A. tequilana has been the biological model of study in the laboratory, having special interest in the fructan metabolic pathway. At the molecular level, there have been reported sequences that translate FT enzymes, specifically, sequences for two isoforms of 1-SST enzyme and two isoforms of 6G-FFT enzyme. Also, there is molecular information available of other members of the PGHF32 in *A. tequilana*, such as additional vacuolar and a cell wall invertases (Cortés-Romero, *et al.,* 2012). More recently, through RNAseq analysis, it was possible to identify new members of PGHF32 in four agave species, including molecular information of several isoforms of FEH enzyme in *A. tequilana*, described for the first time. However, no molecular information could be convincingly detected for 1-FFT and 6-SFT fructosyltransferases (Ávila de Dios, *et al*., 2015).

On the other hand, in the early 2000s, a research group reported the presence of FEH enzymes in the most common plant biological model, *Arabidopsis thaliana* (Arabidopsis) (De Coninck, *et al*., 2005). Arabidopsis is a plant that naturally does not synthesize fructans, meaning that no molecular information of FT enzymes is present in its genome. However, by cDNA purification and heterologous expression analysis in the *Pichia pastoris*

system, it was possible to characterize, at the functional level, three enzymes from Arabidopsis that initially were annotated as putative cell wall invertases.

Cell wall invertase 1 (Atcwinv1) resulted to be a genuine invertase by having sucrose molecules as the preferred substrate, thus confirming its catalytic activity. On the contrary, cell wall invertase 3 and 6 (Atcwinv3, Atcwinv6) were unable to hydrolyze sucrose molecules, showing fructan exohydrolytic activity instead, presenting preference for a levan type fructan (β-(2, 6) linkage) as substrate; Atcwinv6 was also capable of degrading inulin type fructan (β-(2, 1) linkage). Nevertheless, no clear evidence about the possible biological role that these FEH enzymes could be playing in Arabidopsis has been obtained.

In the present research, an Agave-Arabidopsis system was implemented in order to transfer molecular information of FT Atq1SST-1 isoform from *A. tequilana* to Arabidopsis Col. 0 by constructing an expression vector (i. e., pB7WG2D-*Atq1SST-1*) which cDNA was driven under a constitutive promoter (i. e., cauliflower mosaic virus p35S) to transform Arabidopsis through the floral dip method. Having generated several transgenic lines, quantitative gene expression of the transgene (*Atq1SST-1*) as well as the endogenous FEHs (At*cwinv3/6-FEH* and *Atcwinv6/6, 1-FEH*) was carried out, identifying the possibility that the ectopic expression of *Atq1SST-1* could be down regulating *FEH* transcripts in the transgenic lines due to lower expression patterns than those achieved in wild type (WT) Arabidopsis. At the same time, quantitative gene expression of endogenous *FEH* in WT Arabidopsis was carried out by using several tissues and development stages in order to establish the gene expression pattern of these enzymes. Finally, the carbohydrate profile was analyzed both in WT and transformed Arabidopsis, confirming that WT Arabidopsis does not accumulate fructan molecules. However, it was not possible to identify fructan molecules with a degree of polymerization of 3 (corresponding to 1-kestose) in any of the transgenic lines analyzed, which was the expected metabolite to be synthesized by the FT enzyme activity. However, other changes in carbohydrate metabolism were observed.

Background

The *Agave* genus

The *Agave* genus harbours perennial monocarpic plants whose morphology is characterized by rosette like leaves attached to a short stem (also known as “piña” in Mexico). The leaves are commonly large, thick and fibrous. The leaf form may be either linear or lanceolate and usually has small sharp teeth at the lateral edges accompanied with a terminal rigid sharp spine (Figure 1). The environment where these plants normally grow comprises arid or semiarid regions defined by growth-detrimental conditions, such as low water availability and extreme temperatures. Agave plants can tolerate temperatures higher than 40 °C or below 0 °C and most of them are found, besides dry regions, in mountainous or forested areas and coastal beaches (Irish and Irish, 2000).

FIGURE 1. Typical plant morphology of the *Agave* genus. Characteristic leaf morphology (which may vary among species), floral tissues and fruit anatomy are also shown (figure taken from: https://www.freundevonfreunden.com/journal/mezcal/).

The span of time required to reach physiological maturity in agave plants can vary between species, ranging from 10 to 25 years to 4 to 5 years. Once the maturity of the plant has been reached, the reproductive phase starts with the development of the inflorescence. Gentry (1982), divided the *Agave* genus into two subgenera, based on floral characteristics: subgenera *Agave* and subgenera *Littae*. The *Agave* subgenera (Figure 2, A) comprises plants that develop a paniculate inflorescence while *Littae* subgenera (Figure 2, B) includes plants that develop a spicate inflorescence (Irish and Irish, 2000).

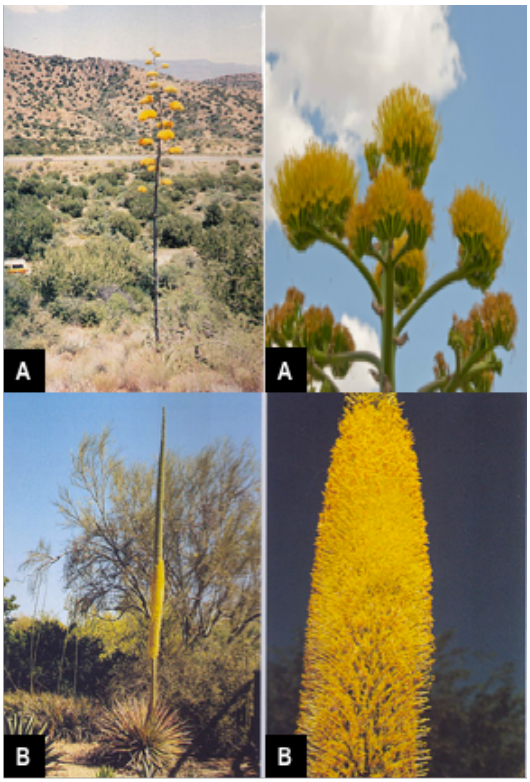


FIGURE 2. Characteristic inflorescence morphology and floral development of **A.** Subgenera *Agave* and **B.** Subgenera *Littae* (image taken and modified from: http://www.agavaceae.com/agavaceae.com/botanik/pflanzen/botanzeige_scan_en.asp?gnr=110&cat=&scan=110-6).

Agave reproduction can be either sexual or asexual. Sexual reproduction can take place once the pollination of the flowers mediated mainly by birds, bats, insects, etc. has occurred. The fruit of agave plants consists of a capsule with three chambers which seeds are located inside each of them. The seeds are flat and black, and a single agave plant

(specifically *A. tequilana*) can produce in average up to 55,000 of them. However, it has been reported that only a low fraction of the seeds is viable for germination (García-Mendoza, 2002; Escobar-Guzmán, *et al*., 2008).

On the other hand, asexual reproduction is accomplished through clones generated either in the basal zone or aerial zone of the mother plant. Clones or suckers (also known as "hijuelos", in Mexico) generated at the basal zone are formed from rhizomes. After some time, the offset forms its own root system but normally remains attached to the mother plant. Bulbil development is the second method of asexual propagation. Bulbils are small agave plantlets generated at the bracteoles of the inflorescence (aerial zone of the mother plant) that fall to the ground and start to grow independently. It has been reported that this strategy occurs when sexual reproduction is unsuccessful or interrupted, ensuring the propagation of the plant in the absence of flower pollination (Arizaga and Ezcurra, 1995; Abraham-Juarez, *et al*., 2010).

As mentioned above, the *Agave* genus can be found in a diversity of regions and is endemic to the American continent (Figure 3). Its distribution extends from southern Canada to Northern South America including the Caribbean Islands (García-Mendoza and Galván, 1995). Mexico, United States of America, Cuba and Guatemala are considered as the countries with the highest number of taxa (García-Mendoza, 2007). In the particular case of Mexico, the *Agave* genus can be found in over 75% of the territory. However, the distribution throughout the Mexican territory is not equal, finding a greater number of species are located in arid, dry regions and decreasing its distribution drastically in humid, warm regions of Southern Mexico such as in the states of Tabasco, Campeche and Quintana Roo (García-Mendoza, 2007).

Agave species have been exploited for several uses such as food, beverages, forage, construction, fibers, medicinal, ornamental, domestic and others (ethanol production, cellulose extraction, etc.) (Nava-Cruz, *et al.,* 2015). Although, *A. tequilana* is also being

used in Australia as a source of biofuel with a high energetic value due to the considerable biomass production of the plant (Holtum, *et al.,* 2011).

FIGURE 3. Geographical distribution of the *Agave* genus (García-Mendoza and Galván, 1995).

Agave tequilana

This agave species was first described and classified in 1902 by the French botanist Frederic Albert Constantine Weber. *A. tequilana* is a perennial monocarpic species and can be considered as an example of a large agave species, having a height range between 1.2 and 1.8 m, leaves between 8-13 cm in width (the widest part of the leaf is located in the middle) and 89-119 cm in length, presenting a specific glaucous bluish or greyish green color. The teeth present at the edges of the leaves are small and either lineal or slightly

curved, separated from each other by 3-6 mm. The terminal spine is short (approximately 1-2 cm long) with a typical dark brown color (Figure 4). The inflorescence is not proportional to the rosette size since it can reach a height of between 5 to 6 m. It is highly branched with 20-25 large umbels where eventually flowers will develop. The taxonomic classification of *A. tequilana* is established as follows:

TABLE 1. Taxonomic classification of *A. tequilana.*

Superkingdom	**Eukaryota**
Kingdom	Viridiplantae
Phylum	Streptophyta
Division	Angiospermae
Class	Mesangiospermae
Order	Asparagales
Family	Asparagaceae
Subfamily	Agavoideae
Genus	*Agave*
Subgenus	*Agave*
Group	*Rigidae*
Species	*Agave tequilana*

The geographical distribution of *A. tequilana* is concentrated in five states within Mexico (i. e., Guanajuato, Jalisco, Michoacán, Nayarit and Tamaulipas) designated under the appellation of origin for the production of tequila (Figure 4). *A. tequilana* var. azul is an important crop in Mexico, because these plants are used as a unique raw material for tequila production where the fructan storing stems (piñas) are exploited to extract sugars for fermentation and distillation of ethanol, as described by the NOM-006-SCFI-2012 in the "Diario Oficial de la Federación".

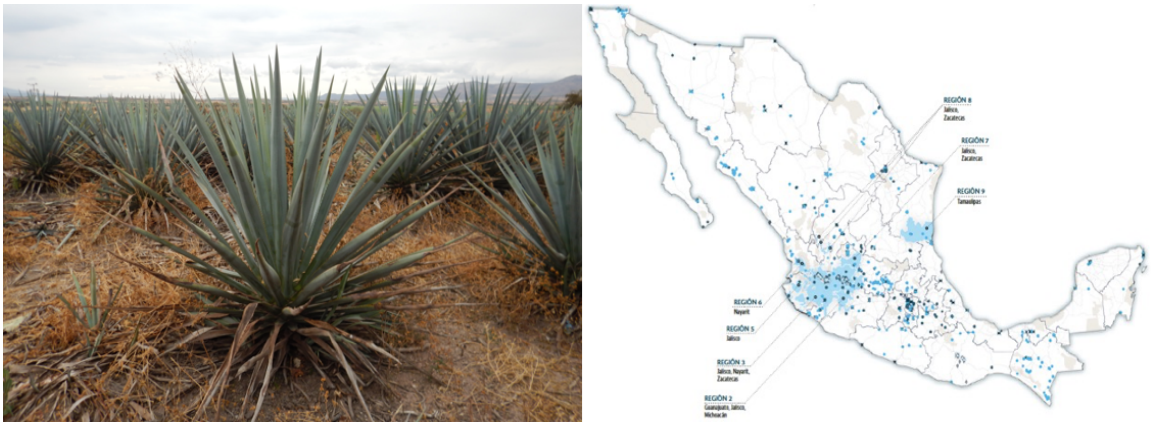


FIGURE 4. *A. tequilana* morphology and geographical distribution in the Mexican territory. Indicated by blue background are the states of Jalisco, Nayarit, Tamaulipas, Michoacán and Guanajuato (figure modified and taken from Agave tequilero y mezcalero Mexicano, 2017, Secretaría de Agricultura Ganadería, Desarrollo Rural, Pesca y Alimentación, SAGARPA).

Crassulacean Acid Metabolism (CAM)

Crassulacean Acid Metabolism (CAM) is a photosynthetic adaptive process that arose in multiple independent evolutionary events. Like C4 metabolism, CAM metabolism uses C4 organic acids during carbon fixation to synthetize intermediate molecules throughout the pathway to finally produce carbohydrates. The characteristics that distinguish CAM from C3 and C4 metabolisms are: nocturnal CO_2 uptake and inverse stomatal behavior (stomata stay closed during the day and open at night). The CAM pathway can be simplified in four steps (Borland, *et al*., 2014, Figure 5):

1. Nocturnal CO_2 uptake starts once stomata are opened. CO_2 assimilation takes place in the cytosol by catalytic activity of phosphoenolpyruvate carboxylase (PEPC), producing oxaloacetate, which is subsequently reduced to malate by malate dehydrogenase to finally be stored in the vacuole as malic acid.
2. Transition between night carboxylation by PEPC to daylight carboxylation by Ribulose bisphosphate carboxylase/oxidase (Rubisco) begins, as photosynthetically active radiation increases in the morning.

3. In daylight, stomata close and mobilization of malic acid from the vacuole to the chloroplast takes place. Malic acid decarboxylation releases three-carbon molecules, used as the carbon source for carbohydrate synthesis, and CO_2 which is re-assimilated by Rubisco regenerating carbohydrates via the Calvin-Benson cycle.

At the end of the day, CO2 assimilation by Rubisco starts to slow-down as organic acid begins to decrease, and the process is repeated.

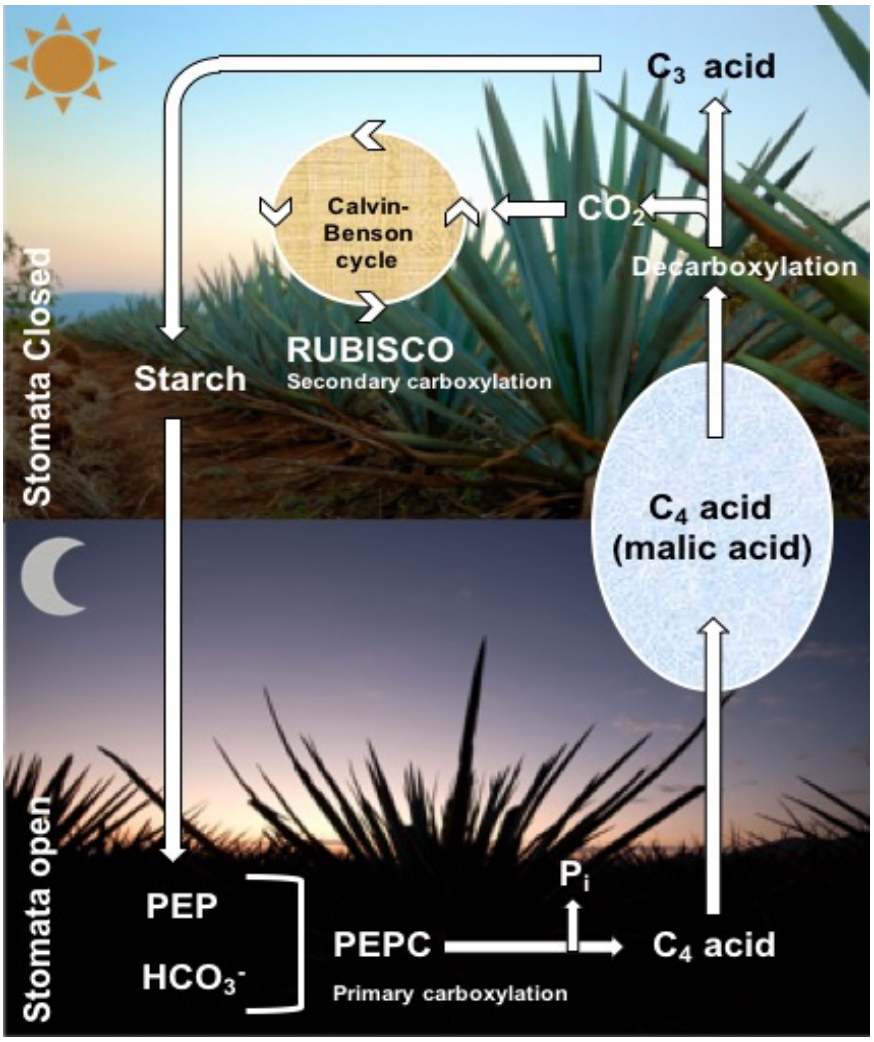


FIGURE 5. Schematic representation of the crassulacean acid metabolism (CAM) metabolism (based on Borland, *et al*., 2014).

Fructans

Fructans are fructose polymers with a glucose residue remaining from the sucrose precursor molecules, whose degree of polymerization (DP) is increased by the sequential addition of fructose molecules catalyzed by the activity of several fructosyltransferases (FTs). These polymers are synthesized by both prokaryotes and eukaryotes (fungi and plants). Focusing on plants, the majority of them synthesize and store carbohydrates in the form of starch or sucrose molecules. However, 15% of angiosperms belonging to families of both monocotyledons and dicotyledons synthesize fructans as the main storage carbohydrate (Hendry, 1993).

There is a high diversity of fructan structures, differing by the DP, presence or absence of branches, the type of linkage between fructose moieties (either β-(2, 1) or β-(2, 6) linkages) and the glucose residue position (internal or terminal). Usually, fructans are classified as follows (Figure 6):

1. **Inulin type fructan** has a linear chain structure in which fructose moieties are linked exclusively by β-(2, 1) glycosidic linkages, presenting a terminal glucose residue. This type of fructan can be found in the Asteraceae family (dicotyledons) (Van den Ende, 2013) in plants such as chicory (*Cichorium intybus*), Jesusalem artichoke (*Helianthus tuberosus*) and *Cynara scolymus*. The simplest inulin type fructan is the trisaccharide 1-kestose.

2. **Levan type fructans** have a linear chain structure with a final glucose residue with fructose moieties linked by β-(2, 6) glycosidic linkages. These types of polymers are mainly synthesized by bacteria such as *Erwinia* and *Pseudomonas*. However, levans (known as phleins) can also be synthesized by plants including some members of the Poaeceae family such as perennial forage grasses like *Poa secunda*

(Wei, et al., 2002) and *Dactylis glomerata* (Bonnett, *et al.,* 1997). The simplest levan type fructan is the trisaccharide 6-kestose.

3. **Graminan type fructan** is characterized by the presence of branches along the chain, whose fructose moieties are linked by either β-(2, 1) or β-(2, 6) glycosidic linkages within the polymer. This type of fructan is mainly predominant in cereals such as wheat (*Triricum aestivum*) (Nilson and Dahlqvist, 1986; Verspreet, *et al*., 2013) and barley (*Hordeum vulgare*) (Sprenger, *et al*., 1995; Duchateau, *et al.,* 2005). The simplest graminan type fructan is the tetrasaccharide bifurcose.

4. **Neoseries type fructan** is synthesized by attaching the first fructose moiety in the sixth carbon from the glucose and whose elongation of the polymer can take place either from the glucose or fructose residues of the initial sucrose molecule. Thus, neoseries type fructan can be divided into inulin neoseries or levan neoseries type fructan, depending on the type of linkage generated to join the subsequent fructose moieties (e. g., β-(2, 1) or β-(2, 6) respectively). Inulin neoseries type fructans can be found in the Alliaceae and Asparagaceae families (Ritsema, *et al.,* 2003; Fujishima, *et al*., 2005; López, *et al.,* 2003) whereas levan neoseries type fructans can be found in the Poaceae family (Livingston, *et al.,* 1993; Lüscher, *et al.,* 1995; Pavis, *et al.,* 2001).

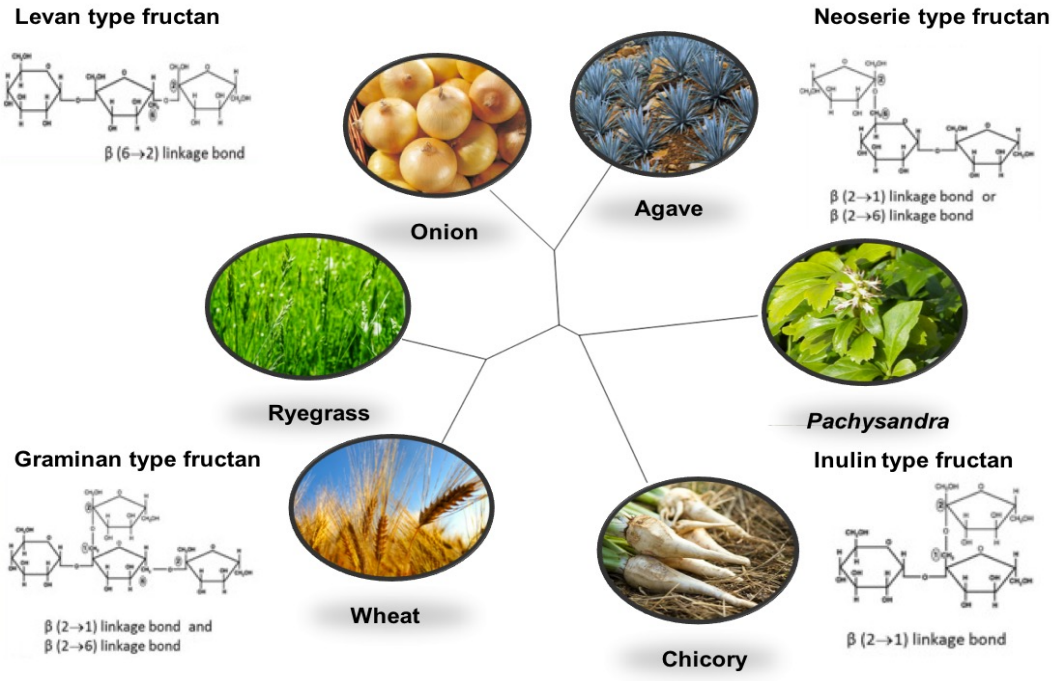


FIGURE 6. Chemical structures of the different fructan types and the common plant species where they can be found (figure based on and modified from Van den Ende, *et al.*, 2013 and Cimini, *et al*., 2015).

Plant Glycoside Hydrolase Family 32 (PGHF32)

Glycoside hydrolases are enzymes whose catalytic activity causes the hydrolysis of the glycoside bonds between two carbohydrates or between a carbohydrate and a non-carbohydrate moiety. These types of enzymes are divided into several families based on amino acid sequence similarities. Currently, there are 135 families of glycoside hydrolases (Cimini, *et al.,* 2016). More specifically, Plant Glycoside Hydrolase Family 32 (PGHF32) includes different types of enzymes involved in sucrose breakdown such as cell wall invertases (Cwinv) and vacuolar invertases (Vinv), enzymes exhibiting transglycosylating activities known as fructosyltransferases (FT) and enzymes involved in fructan breakdown known as fructan exohydrolases (FEH). In addition to the plant enzymes, GHF32 also includes inulinases, exoinulinases and levanases from fungi (www.cazy.org/GH32.html). GHF32 is incorporated with GHF68 in clan GH-J. GHF68 (www.cazy.org/GH68.html) includes invertases, levansucrases and inulosucrases from prokaryotes. These families are

united into a single clan due to similarities in their three-dimensional protein structure. Although, both families share only the N-terminal β-propeller domain (Van den Ende, 2013) and less than 15% of similarity at the amino acid sequence level (Naumoff, 2001).

The three-dimensional structure of GHF32 enzymes has been determined for several organisms, showing a commonly conserved fold. The core of the structure is characterized by an N-terminal β-propeller domain and the C-terminal by two antiparallel six-stranded β-sheets. The β-propeller domain is characterized by the presence of five-fold repeated blades which each contain four antiparallel β-strands around a central cavity (Figure 7). The catalytic site of these enzymes is located within the β-propeller domain and, at the molecular level, is characterized by the presence of acidic residues in three highly conserved motifs (also known as the catalytic triad). The first motif WMN**D**PNG contains an aspartic acid acting as the catalytic nucleophile, the **E**C motif contains a glutamic acid acting as the general acid/base and finally, the R**D**P motif contains a second aspartic acid that acts as a transition-state stabilizer in the catalytic reaction.

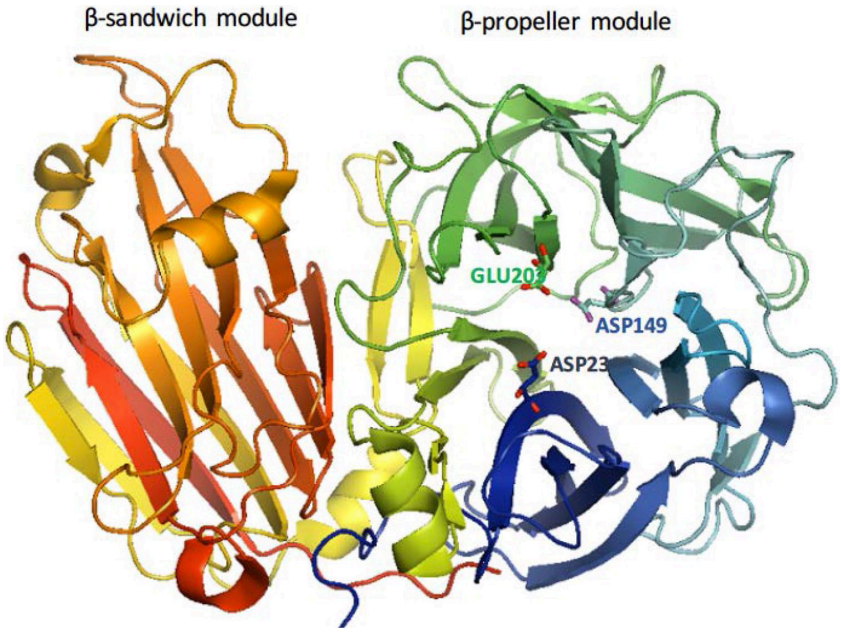


FIGURE 7. Conserved three-dimensional protein structure of the Glycoside Hydrolase Family 32 (GHF32) members (figure taken from Cimini, *et al.*, 2016).

To date, the three-dimensional structures of GHF32 enzymes from plants that have been analyzed by crystallography include: a cell wall invertase (Atcwinv1) from Arabidopsis (Verhaest, *et al*., 2006), a dual function fructosyltransferase (6-SST/6-SFT) from *Pachysandra terminalis* (Lammens, *et al*., 2012) and a fructan exohydrolase (1-FEH IIa) from *C. intybus* (Verhaest, *et al*., 2007) (Figure 8).

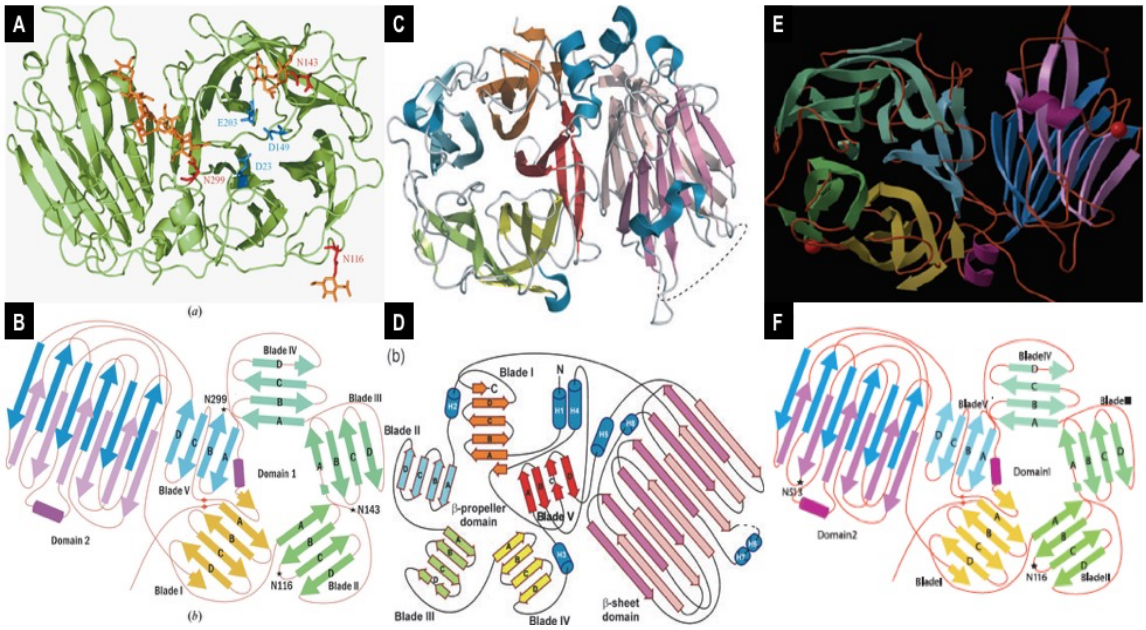


FIGURE 8. Three-dimensional structures described and reported by protein crystallization for a cell wall invertase (Atcwinv1) from Arabidopsis (**A, B**; Verhaest, *et al*., 2006), a dual function fructosyltransferase from *Pachysandra terminalis* (6-SST/6-SFT) (**C, D**; Lammens, *et al*., 2012) and a fructan exohydrolase (1-FEH IIa) from *C. intybus* (**E, F**; Verhaest, *et al*., 2005).

Fructan biosynthetic pathway

Edelman and Jefford (1968) proposed the first biosynthetic inulin pathway from *H. tuberosus*, explaining that for inulin type fructan synthesis, the activities of two fructosyltransferases, i. e., sucrose: sucrose 1-fructosyltransferase (1-SST) and fructan: fructan 1-fructosyltransferase (1-FFT), are necessary. Conversely, hydrolase activity is needed for fructan degradation. Additionally, they suggested that the tonoplast (i. e., the

vacuole delimiting membrane) is the functional localization where both reactions (synthesis and breakdown) take place.

In dicotyledons, two enzymes are responsible for fructan synthesis, as predicted by the classical Edelman and Jefford model. The catalytic activities of 1-SST and 1-FFT are sufficient to synthesize linear inulins comprising 10 to 200 fructose moieties depending on the plant species (Van den Ende, *et al*., 2006). In contrast to dicotyledonous species, monocotyledons require 3 - 4 fructosyltransferases in order to synthesize a wide diversity of fructans, differing mainly in branching and glycosidic linkages. Fructan biosynthesis in monocotyledons (Vijn and Smeekens, 1999; Banguela and Hernandez, 2006; Yildiz, 2010; Cimini, *et al.,* 2015) can be described as follows (Figure 9):

1. The synthesis of fructan molecules starts by the catalytic activity of sucrose: sucrose 1-fructosyltransferase (1-SST, EC 2.4.1.99) using two sucrose molecules, where one of them donates the fructosyl moiety, which is linked to the second sucrose molecule through a β-(2, 1) linkage forming the trisaccharide 1-kestose.
2. Fructo-oligosaccharides (FOS) can be elongated by the catalytic activity of fructan: fructan 1-fructosyltransferase (1-FFT, EC 2.4.1.100) using either sucrose or fructan molecules (such as 1-kestose or fructans with a higher DP) as the fructosyl donors, linking the fructosyl unit to another fructo-oligosaccharides in order to generate a molecule with a higher DP.
3. When both 1-kestose and sucrose are available, sucrose: fructan 6-fructosyltransferase (6-SFT, EC 2.4.1.10) synthesizes the graminan type fructans. Sucrose molecules act as fructosyl donors and 1-kestose acts as the acceptor. In this case, the fructosyl unit is linked by a β-(2, 6) linkage. However, this is not the only catalytic function performed by this enzyme. When sucrose molecules are the sole substrate, 6-SFT synthesizes levan type fructans, where a sucrose molecule donates the fructosyl unit to another sucrose acceptor molecule linking the fructosyl unit by a β-(2, 6) glycosidic linkage.

4. Finally, the last group of fructans, the neoseries type, is synthesized by the activity of fructan: fructan 6G-fructosyltransferase (6G-FFT, EC 2.4.1.243). This enzyme requires the fructosyl unit from a sucrose molecule donor, that will be joined by a glycosidic linkage to a 1-kestose molecule. The fructosyl unit is linked by a β-(2, 6) glycosidic bond to the sixth carbon of the glucose residue to form neokestose. Therefore, the neokestose molecule can be further elongated either by β-(2, 1) or β-(2, 6) glycosidic linkages, synthesizing inulin or levan neoseries type fructans respectively.

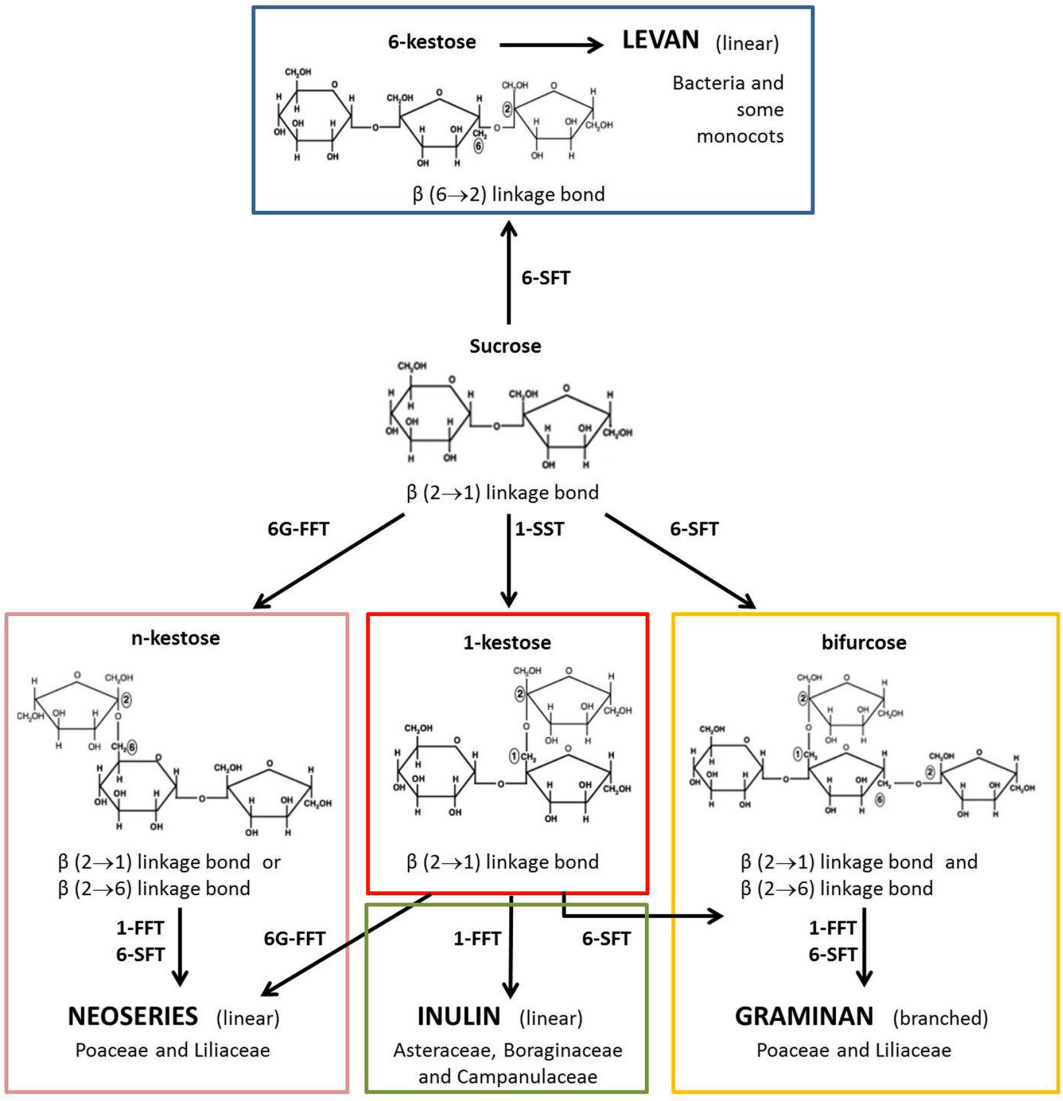


FIGURE 9. Fructan synthetic pathway and the primary chemical structures of the different fructan types produced by the catalytic activity of the diverse fructosyltransferases (figure taken from Cimini, *et al.*, 2015).

Fructan breakdown

Mobilization of fructan reserves usually begins when the plant is under active growth. The enzymes responsible for fructan hydrolysis are known as fructan exohydrolases (FEH), releasing terminal fructosyl units using water molecules as an acceptor (Valluru and Van den Ende, 2008). These FEH enzymes are classified on the basis of the nature of the bonds they hydrolyze. Fructan 1-exohydrolase (1-FEH, EC 3.2.1.153), carries out the hydrolysis of β-(2, 1) glycosidic linkages mainly from inulin type fructans whereas fructan 6-exohydrolase (6-FEH, EC 3.2.1.154) carries out the hydrolysis of β-(2, 6) glycosidic linkages from levan type fructans (Figure 10). FEH enzymes are unable to degrade sucrose molecules and, as a matter of fact, sucrose molecules act as a competitive inhibitor of the FEH enzymes (Verhaest, *et al.,* 2007). Fructan endohydrolases are absent in plants (Yildiz, 2010; Versluys, *et al.,* 2018).

Nevertheless, some FEH variants have been reported, e.g. a Fructan 6, 1-exohydrolase, (6,1-FEH) from wheat (*T. aestivum*) capable of hydrolyzing bifurcose (a graminan type fructan). This enzyme also presented affinity (albeit at a lower level) for β-(2, 6) glycosidic linkages (levan type fructan) and it was suggested that this enzyme plays a crucial role during cold stress (Kawakami, *et al.,* 2005). Lastly, Van den Ende and colleagues (2005) reported two enzymes capable of hydrolyzing 6-kestose molecules as sole substrates in wheat (*T. aestivum)* identifying sucrose and fructose as the main products from the catalytic activity of the enzymes. These enzymes were denominated 6-kestose exohydrolases (6-KEH w1 and 6-KEH w2).

Invertases

Invertases (EC. 3.2.1.26) are enzymes that catalyze the irreversible hydrolysis of sucrose molecules, generating as a result the glucose and fructose monosaccharides (Figure 10). Based on subcellular localization and conditions for optimum catalytic activity, such as

pH, solubility and isoelectric point, three kinds of invertases could be identified: vacuolar, cell wall and cytosolic invertases (Tymowska-Lananne and Kreis, 1998). Sucrose, glucose and fructose may act as signaling molecules and have been directly related to plant growth, flowering regulation, development of storage organs, etc. (Tauzin and Giardina, 2014). Therefore, invertases are important enzymes, due to their roles in gene regulation and plant development (Koch, 1996).

As mentioned previously, vacuolar and cell wall invertases are members of Plant Glycoside Hydrolase Family 32 (PGHF32). Both invertases share certain biochemical characteristics, such as the optimum pH of catalytic activity, which is between 4.5 and 5.0 (therefore their denomination as “acid invertases”) and both enzyme types are strongly inhibited by the products of the hydrolysis reaction.

Cell wall invertases (cwinv) are key enzymes in carbohydrate metabolism because they regulate sucrose hydrolysis and are most active in apoplastic sinks. On the other hand, they are also involved in pollen and seed development. Cwinvs are also thought to participate in the wound response, triggering plant defenses by inducing the expression of defense related genes when the plant has been infected by a pathogen (bacteria, fungi, viruses, nematodes, etc.) (Tauzin and Giardina, 2014) or damaged by an herbivore.

On the other hand, vacuolar invertases (vinv) are enzymes that regulate the sucrose-hexose/ osmoregulation ratio in vacuoles and are normally most active in growing and/ or expanding tissues such as root tips, stem internodes and developing leaves (Weschke, *et al*., 2003). Vinvs, unlike cell wall invertases, are not thought to be directly involved in plant defense response (Yildiz, 2010; Tauzin and Giardina, 2014).

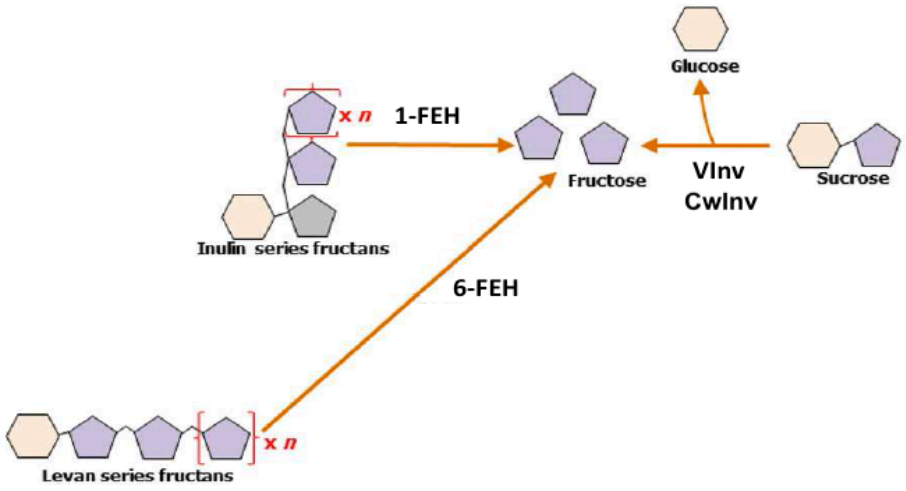


FIGURE 10. Fructan and sucrose hydrolysis pathways by fructan exohydrolases (FEH) and vacuolar/ cell wall invertases (Vinv, Cwinv) enzymes (figure taken and modified from Tarkowski and Van den Ende, 2015).

The role of fructans in plants

Fructan biosynthesis in plants is closely related to the production of fructan polymers which serve as the major reserve carbohydrates in 15% of angiosperms, acting similarly to starch or sucrose, thereby, being accumulated when carbon molecules exceed the typical demand and mobilized when energy is required. They are synthesized, stored and hydrolyzed in the vacuole (Vijn and Smeekens, 1999; Banguela and Hernandez, 2006) and might modulate overall sucrose concentration in the vacuole, lowering the concentration of sucrose in order to prevent sugar induced inhibition of photosynthesis (Vijn and Smeekens, 1999). Fructans have also been associated with petal expansion and flower development in *Hemerocallis* (Bieleski, 1993), *Campanula rapunculoides* (Vergauwen, *et al*., 2000) and recently in *A. tequilana* (Ávila de Dios, *et al.,* 2015).

Moreover, fructans might play an important role as protective molecules against abiotic stress conditions, such as drought and freezing. This is because the fructan biosynthetic pathway and photosynthesis can be performed normally at low temperatures whereas starch biosynthesis drops drastically in temperatures below 10 °C (Pollock, 1986). Water

molecules play a crucial role in plasma membrane stability, by establishing hydrogen bonding bridges among lipid head groups. Under abiotic stress, cells experiment water molecule loss from the bilayer membrane, causing the cells to shrink and cell membrane fusion then takes place, leading to serious cell injury. It has been suggested, that sugars such as hexoses can substitute water molecules during drying (Valluru and Van den Ende, 2008; Livingston, *et al.,* 2009), preventing plasma membrane fusion by interacting directly through hydrogen bonds with the polar head-groups of phospholipids.

Fructans are the only polysaccharide that can interact with lipid head-groups, being capable of inserting between lipid head-groups and stabilizing lipid bilayers (Valluru and Van den Ende, 2008), due to the presence of flexible furanose rings, in contrast to other polysaccharides composed of more rigid pyranose rings (such as galactose, glucose or mannose). Superior membrane protection is provided by a mixture of high and low DP fructans, as shown by differences in membrane protection related with fructan DP in different cold resistant cereals (Suzuki and Nass, 1988; Livingston, *et al.,* 2007). Additionally, the presence of FEH enzymes and fructans in the apoplast (Livingston and Henson, 1998; Van den Ende, *et al*., 2005), raises the possibility that apoplastic fructans could be hydrolyzed by FEH enzymes, generating a mixture of fructan oligosaccharides, sucrose and fructose that can participate in plasma membrane stabilization. Studies carried out in oat, determined that hexoses from fructan breakdown could lower the freezing point of water by one degree (Pollock, *et al.*, 1988; Livingston and Henson, 1998).

Small sugar molecules as well as carbohydrates with a higher DP, not only can act as possible source of food for the pathogens but also, they can act as signaling molecules that can induce plant defense responses. In the last years, it has been proposed that sugar molecules could act as damage-associated molecular patterns (DAMPs) when present in the apoplast in response to an abiotic stress event. Similarly, they have also been proposed to function as microbe-associated molecular patterns (MAMPs), which are associated to the synthesis of levan type fructans, which increase pathogen virulence

through biofilm formation and calcium chelation to prevent plant defense. However, these levans can be hydrolyzed by FEH enzymes, located in the apoplast, and dispersed throughout the apoplast to trigger defense-related responses that could be triggered in the plasma membrane by unknown receptors.

Versluys and colleagues (2017) suggested a model explaining the possible way that fructans could act as DAMPs and/ or MAMPs. In animals, during biotic stress, it is thought that microbial fructans could be perceived by Toll-like receptors (TOR) 2 and 4 in the host gut systems, being recognized as MAMPs molecules, thus inducing immune responses. In plants, when a phytopathogen infects a host, it has been hypothetically proposed that microbial fructan molecules could be sensed by an unknown receptor, and therefore recognized as MAMPs. On the other hand, fructans may also act as DAMPs by their recognition, probably by the same unknown receptor responsible for activating the MAMP-related responses. During abiotic stress, cell rupture is one of the common results of the stress, leading to the leakage of fructan molecules stored in the vacuole into the apoplast, triggering priming in the neighboring cells and upgrading an immune response after fructan molecule recognition by these unknown receptors. Thus, abiotic stress might enhance pathogenic response against further phytopathogen infections (Figure 11).

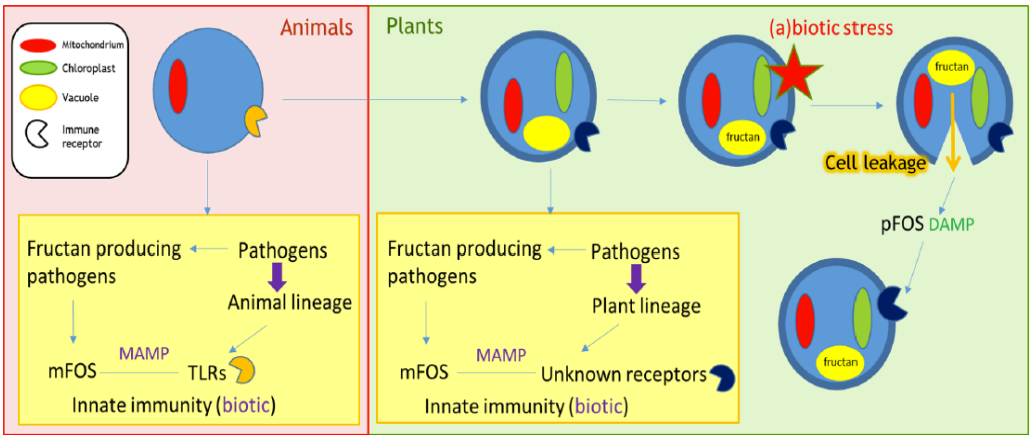


FIGURE 11. Proposed model about the role of fructan molecules as MAMPs in animals and the possible pathway of fructans as MAMPs and DAMPs in plants (Versluys, *et al.*, 2017).

Fructans in *A. tequilana*

López and colleagues (2003), characterized fructan molecules from 8 year old *A. tequilana* plants at the biochemical and structural level. By analyzing the type of glycosidic linkages, it was discovered that fructans synthesized by *A. tequilana* are not only of the inulin type but can be fructans with β-(2, 6) glycosidic linkages. Thus, the presence of fructans with both internal and external glucose residues was reported (Figure 12, A). Likewise, through a comparative analysis in fructan content from several agave species (*A. tequilana, A. angustifolia, A. fourcroydes, A. potatorum* and *A. contala*) and *Dasylirion* spp. grown in different agricultural regions of Mexico (i. e., Jalisco, Guanajuato, Sonora, Oaxaca, Yucatán and Chihuahua), it was possible to determine that fructan content is directly related to both the geographical region and the agave species, reporting that fructan content reaches approximately 60% of total water-soluble carbohydrates in the *A. tequilana* stem (Figure 12, B). The principal fructans present in agave were found to have a terminal glucose (graminan type fructan) (Figure 13, A) or be neoseries type fructans (Mancilla-Margalli, *et al*., 2006).

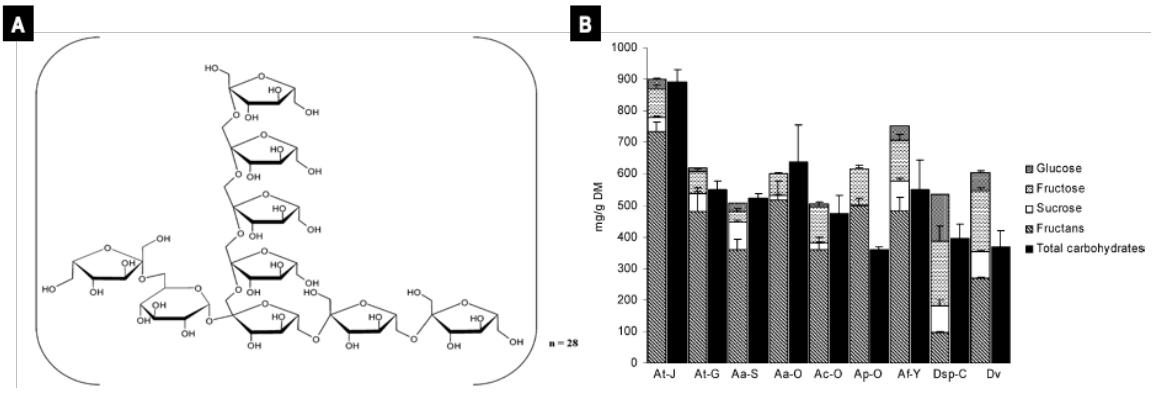


FIGURE 12. A. Chemical structure proposed for fructans synthesized by *A. tequilana* (López, *et al*., 2003) and **B**. Water-soluble carbohydrate content from several agave species grown in different agricultural regions in Mexico (Mancilla-Margalli, *et al*., 2006).

Given their complex, branched structure, a new class of fructans, coined as "Agavins", was proposed by Mancilla-Margalli and colleagues (2006). The Agavins are described as

neoseries type fructans due to the presence of an internal glucose residue, but the main fructose chain may also be branched like graminan-type fructans (Figure 13, B). The authors concluded that the "Agavins" are highly complex fructan molecules, synthesized exclusively by *A. tequilana* plants.

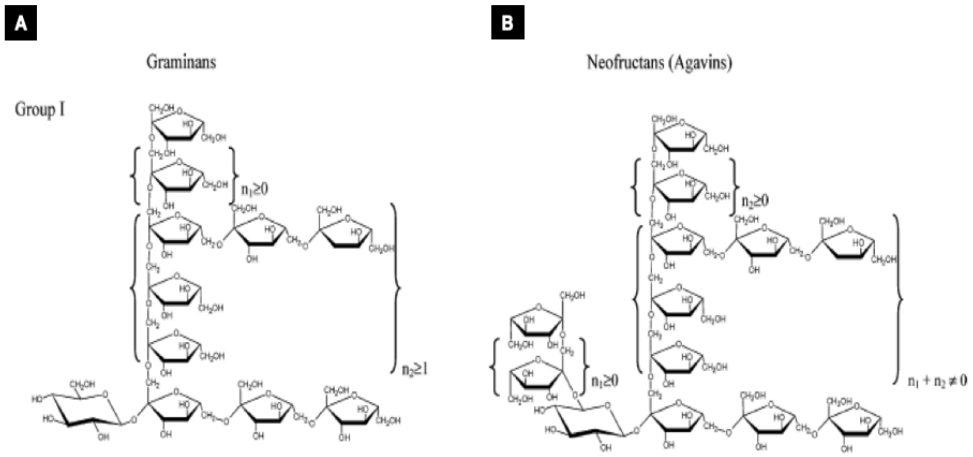


FIGURE 13. Chemical structures of **A.** graminan type fructans and **B**. "Agavin" type fructans proposed to the reference fructans synthesized exclusively by *A. tequilana* plants (Mancilla-Margalli, *et al*., 2006).

Fructan accumulation throughout the *A. tequilana* life cycle is accumulative, which means that in the early life cycle the concentration of fructan molecules tends to be low, increasing throughout the agave development and showing two stages where fructan content decreases, related to asexual propagation through rhizomes and inflorescence development, respectively (Figure 14, A) (Mellado-Mojica, *et al.*, 2009). Likewise, by analyzing the fructan metabolism throughout the *A. tequilana* life cycle (2 to 7 years old plants) it was determined that in the early stages of development of the plant (2- 4 years) the carbohydrate profile is characterized by the presence of fructo-oligosaccharides (FOS) with a low DP while in older plants the fructan profile show presence of complex structures with a high DP, represented mainly by "Agavins" and graminan type fructans (Figure 14, B) (Mellado-Mojica, *et al.,* 2012).

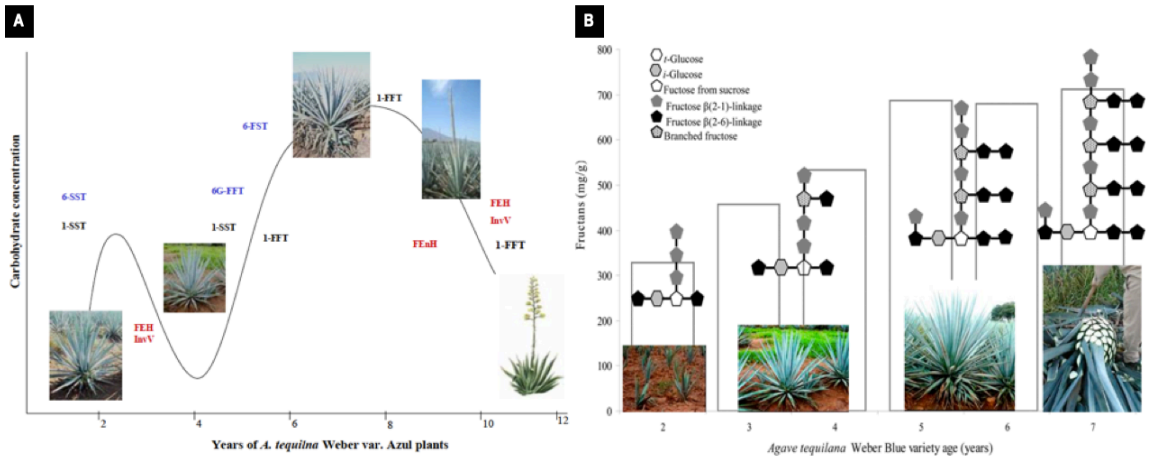


FIGURE 14. A. Theoretical activity of PGH32 throughout *A. tequilana* life cycle and **B**. Fructan content and fructan chemical structures complexity throughout the *A. tequilana* life cycle (figures taken from Mellado-Mojica, *et al.*, 2009 and 2012).

Plant Glycoside Hydrolase Family 32 in *A. tequilana*

The DNA sequence and derived amino acid sequence of enzymes from PGHF32 in *A. tequilana* were first reported in 2007 (Avila-Fernandez, *et al.,* 2007), when the first fructosyltransferase (1-SST) was identified and characterized at the molecular and functional level by heterologous expression in *Pichia pastoris*.

Cortés-Romero and colleagues (2012) described 6 new members of PGHF32 from *A. tequilana* where they identified and characterized, at the molecular level, two isoforms of the 1-SST enzyme (Atq1SST-1 and Atq1SST-2) and for the first time, two isoforms of the 6G-FFT enzyme (Atq6GFFT-1 and Atq6GFFT-2) along with a cell wall invertase (AtqCwinv) and a vacuolar invertase (AtqVinv) (Figure 15). At the functional level, only half of the cDNAs (Atq1SST-2, Atq6GFFT-1 and AtqCwinv) were cloned and expressed heterologously in *P. pastoris,* thereby confirming the catalytic activity of these enzymes from *A. tequilana.* Unexpectedly, AtqCwinv was capable of hydrolyzing 1-kestose molecules, in addition to sucrose, indicating at least a low level of FEH activity. This research did not uncover any cDNAs that convincingly encoded either 6-SFT or FEH type enzymes.

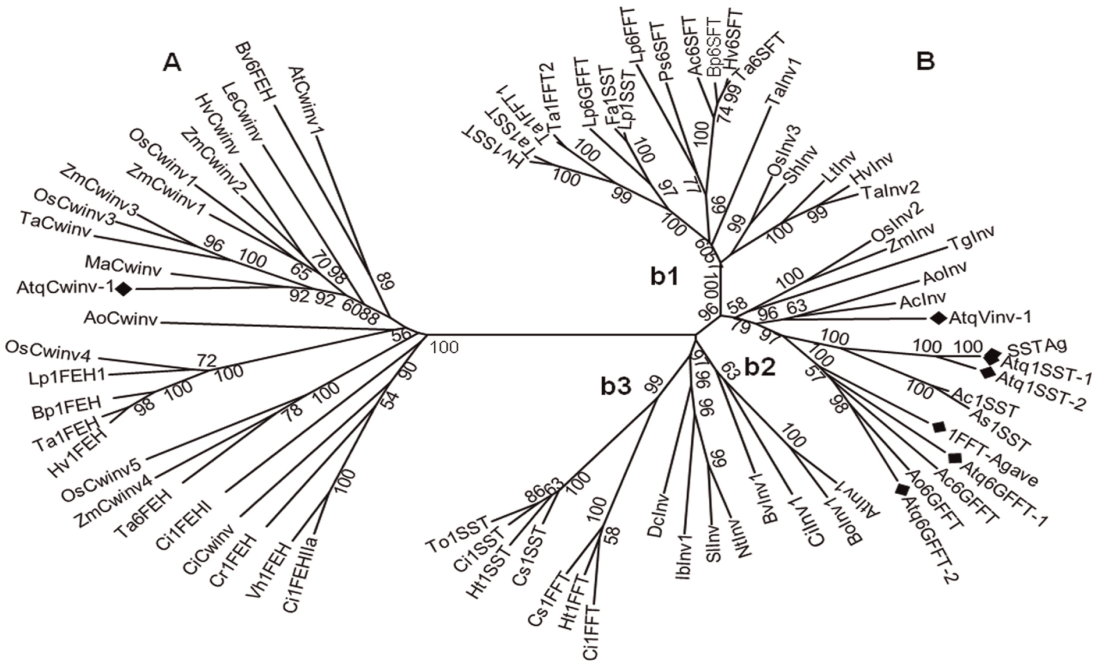


FIGURE 15. Unrooted tree of amino acid sequences of PGHF32 enzymes from dicotyledonous and monocotyledonous plant species. **A.** Clade containing cell wall invertases and fructan exohydrolases and **B.** Clade containing fructosyltransferases and vacuolar invertases. *A. tequilana* enzymes are marked with black diamonds (Cortés-Romero, *et al*., 2012).

Finally, Ávila de Dios and colleagues (2015) reported 31 new full-length cDNA members of PGHF32 from four agave species (*A. tequilana, A. deserti, A. striata* and *A. victoriae-reginae*) (Figure 16). For the first time, it was possible to identify sequences putatively encoding FEH enzymes from all agave species, in addition to PGHF32 members from *A. deserti, A. striata and A. victoriae-reginae* which were reported for the first time. Focusing on *A. tequilana*, it was possible to detect new isoforms for degrading enzymes (invertases and FEH mainly) and by *in silico* expression analysis, it was possible to have a whole perspective of levels and patterns of expression of PGHF32 encoding genes in different tissues and in different agave species. Nevertheless, in all the transcriptome databases analyzed, no sequences for 1-FFT or 6-SFT enzymes were identified.

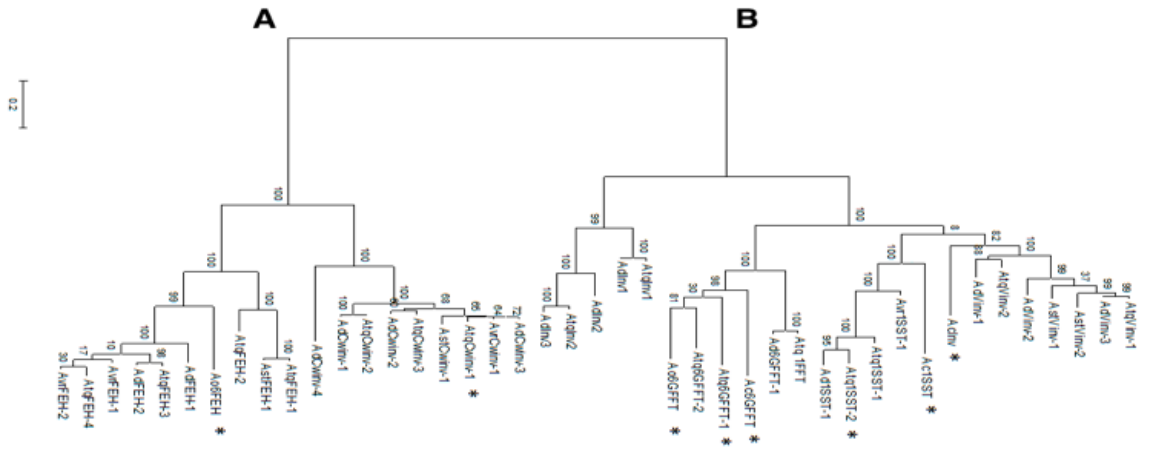


FIGURE 16. Unrooted tree of amino acid sequences of PGHF32 enzymes discovered in agave species and amino acid sequences from close members of the Asparagales order (i. e., *Allium cepa* and *Asparagus officinalis*). **A.** Branch gathering cell wall invertases and fructan exohydrolases and **B.** Branch gathering fructosyltransferases and vacuolar invertases. Those enzymes marked with an asterisk were previously characterized at the molecular and functional level (modified from Ávila de Dios, *et al*., 2015).

Tailor-made fructans

Transfer of genes of interest to model systems and commercial varieties has been possible for over 2 decades and usually has aimed at functional gene characterization or crop improvement for desired traits. Transgenic, non-fructan accumulating plants expressing genes encoding bacterial and plant fructosyltransferases have been successfully developed. According to Banguela and Hernandez (2006), the transgenic production of plant derived fructans has two main goals: the study of the enzyme activities associated with distinct fructan biosynthesis pathways and the creation of novel fructan-producing crops of biological interest.

Levansucrase genes from different bacterial strains have been cloned and expressed under distinct promotors in plant hosts such as tobacco, potato, maize and sugar beet (reviewed by Banguela and Hernandez, 2006; van Arkel, *et al*., 2013). Although bacterial

levansucrase expression was possible in non-fructan accumulating plants, most of the transformed plants displayed aberrant phenotypes such as stunted growth and leaf bleaching, possibly due to cell toxicity caused by bacterial levan accumulation, either in the vacuole or other cellular compartments. Fructan yield was low in most of the transformed plants reported (reviewed by Banguela and Hernandez, 2006; van Arkel, *et al.*, 2013).

Plant fructosyltransferases genes transferred to non-fructan accumulating plants showed neither phenotypic aberration nor fitness alterations and the genes were completely functional. There are reports where each FT enzymes have been cloned, individually or in combination, and have been expressed in non-fructan accumulating plants. In most of the cases, *1-SST* gene expression resulted in the synthesis of 1-kestose (DP3) and nystose (DP4) molecules (Van den Meer, *et al.,* 1998). *6-SFT* gene expression resulted in phlein type fructan synthesis when the enzyme had sucrose as a sole substrate and traces of graminan type fructans were observed in chicory plants due to the presence of endogenous inulin type fructans (Sprenger, *et al.,* 1997). The *6G-FFT* gene needed to be provided with exogenous 1-kestose molecules in order to have a normal catalytic activity, resulting in the formation of inulin neoseries type fructans (Vijn, *et al.,* 1997). By transferring a combination of two enzymes, it could be possible to detect an increase in the DP but also, the fructan profile was found to be similar to the fructan accumulation in the species where the FT genes came from (Van der Meer, *et al*., 1998; Weyens, *et al.,* 2004; Stoop, *et al.,* 2007).

Fructan exohydrolase (FEH) activity in non-fructan accumulating plants.

Based on sequence comparisons, several authors have suggested that an ancestral invertase gene diverged to produce vacuolar and cell wall invertases (Wei and Chatterton, 2001; Banguela and Hernandez, 2006; Van den Ende, 2013). Further divergence led to the

evolution of fructosyltransferases from vacuolar invertases and fructan exohydrolases from cell wall invertases. At the molecular level, it is practically impossible to predict the catalytic function from these enzymes based on amino acid sequence analysis alone, due to highly homologous sequences.

Surprisingly, in some non-fructan accumulating plants such as sugar beet (*Beta vulgaris)* and Arabidopsis*,* fructan exohydrolase (FEH) activity has been reported. In both cases, the FEH enzymes were originally annotated as cell wall invertases. However, by cloning, purifying and characterizing the enzymes, it was possible to demonstrate preferential FEH activity rather than invertase activity. In *B. vulgaris*, the putative FEH showed no enzymatic activity when incubated with sucrose, confirming the lack of invertase activity. On the contrary, catalytic activity was detected when the enzyme was incubated with substrates consisting of β-(2, 6) glycosidic linkages. The authors also tested the presence of fructan molecules in sugar beet by chromatographic techniques, but none could be detected (Van den Ende, *et al*., 2003).

Similarly, the Arabidopsis genome was annotated with 6 putative cell wall invertases (Atcwinv1-6) that showed differential expression patterns. However, De Coninck and colleagues (2005), based on Van den Ende and colleagues research (2003), demonstrated that some cell wall invertases sequences from Arabidopsis were found in different clades within the phylogenetic tree, raising the possibility that some of the putative cell wall invertases could be actually FEH enzymes (Figure 17). Atcwinv 1, 3 and 6 were chosen as the invertases targets for the analyzes due to lack of previous molecular and functional characterization and by heterologous expression in *P. pastoris*, Atcwinv1 was confirmed to catalyze invertase activity. However, the enzyme was also capable of degrading fructo-oligosaccharides but at a lower rate. Atcwinv3 was inactive in the presence of sucrose molecules and could degrade fructan molecules, preferring those linked by β-(2, 6) glycosidic linkages, confirming FEH catalytic activity. Finally, Atcwinv6 was shown to have some invertase activity albeit at a very low rate but showed highest activity in fructan-

oligosaccharide degradation. Atcwinv6 had a higher preference for levan type fructans degradation but could also degrade inulin type fructans. Thus, the presence and activity of FEH enzymes in Arabidopsis was clearly demonstrated, suggesting that Atcwinv3 could be classified as a 6-FEH and Atcwinv6 could be described as a (6, 1)-FEH due to the capability to degrade β-(2, 6) and β-(2, 1) glycosidic linkages respectively.

The presence and function of FEH enzymes in non-fructan accumulation plants is not straightforward. However, the putative role of these enzymes has been hypothesized as participants in the degradation of small amounts of endogenous and/ or exogenous molecules, or prevention of microbial infections. FEHs may also have a detoxification role by preventing fructan accumulation after a microbial infection has occurred (De Coninck, *et al.*, 2005) of fructan signaling (Versluys, *et al.,* 2017).

Based on this information, the main objectives of this work were to express the *A. tequilana Atq1SST-1* cDNA in Arabidopsis*,* determine the presence (if any) of low DP fructans in transformed lines and by last, determine whether *Atq1SST-1* expression affected the pattern of expression of endogenous *FEH* genes (*Atcwinv3* and *Atcwinv6*).

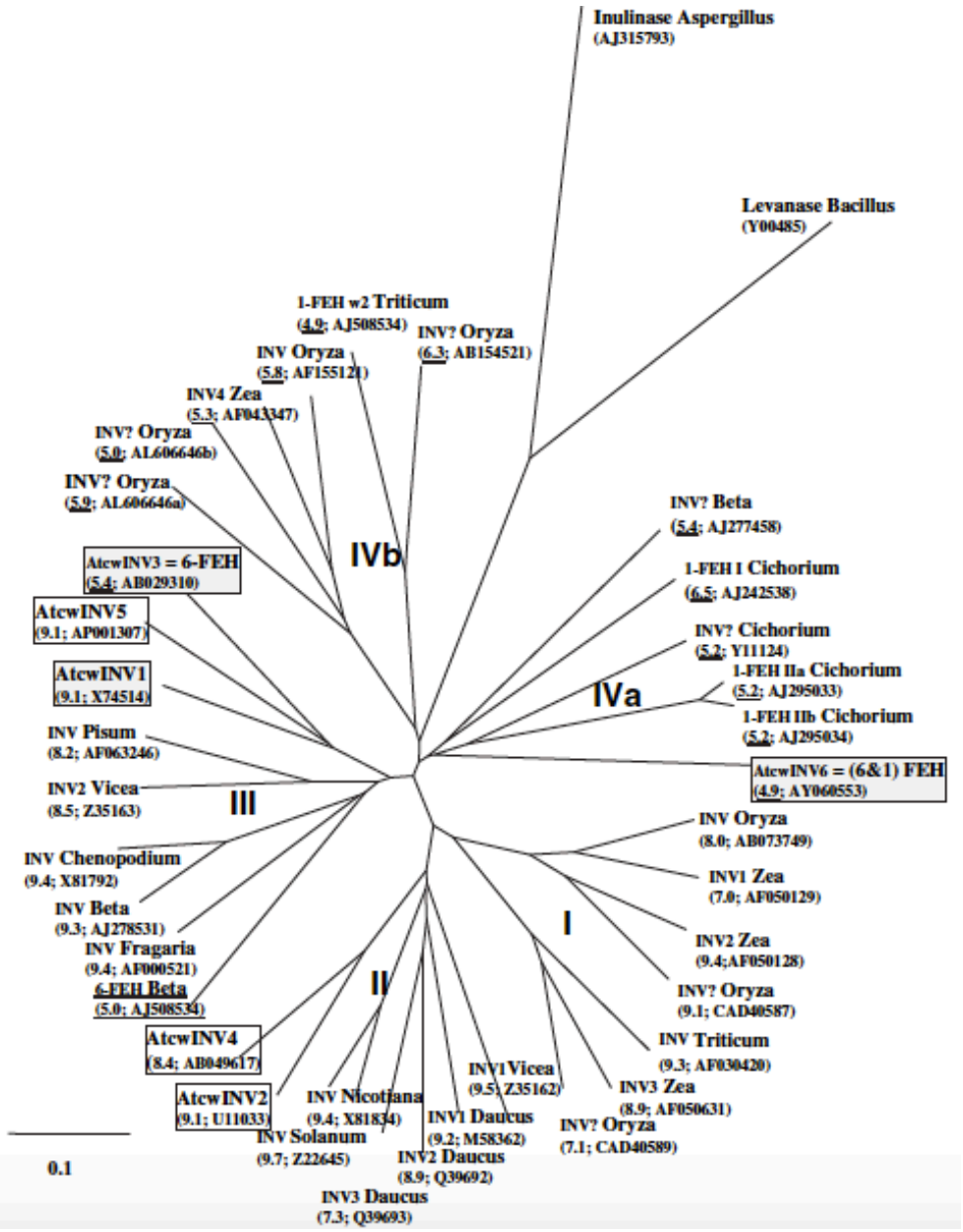


FIGURE 17. Unrooted tree containing amino acid sequences derived from cDNA of cell wall invertases (Cwinvs) and fructan exohydrolases (FEHs) from dicotyledonous and monocotyledonous plants. Arabidopsis cell wall isoforms are located in different groups along with another cell wall invertases or FEHs in black squares (De Coninck, *et al*., 2005).

Hypothesis

The expression of the *A. tequilana Atq1SST-1* cDNA in Arabidopsis will lead to the synthesis of low DP fructans, thereby affecting the expression pattern of the endogenous *FEH* genes (*Atcwinv3* and *Atcwinv6*).

Main objective

To achieve the synthesis of low DP fructans in Arabidopsis by transferring and expressing the *A. tequilana Atq1SST-1* cDNA and determine the putative effects on the expression patterns of endogenous *FEH* (*Atcwinv3* and *Atcwinv6*) genes.

Specific objectives

1. To clone the *Atq1SST-1* cDNA in the Gateway/ *Agrobacterium tumefaciens* vector system.
2. To obtain transgenic Arabidopsis lines expressing *Atq1SST-1* isoform and confirm expression by qRT-PCR.
3. To determine the presence of low DP fructans in transgenic Arabidopsis lines expressing *Atq1SST-1* by thin layer chromatography (TLC).
4. To characterize the expression levels of endogenous *FEH* genes (*Atcwinv3* and *Atcwinv6*) in WT Arabidopsis.
5. To characterize the expression levels of endogenous *FEH* genes (*Atcwinv3* and *Atcwinv6*) in transgenic Arabidopsis expressing the *Atq1SST-1* transgene.

Materials and methods

A. tequilana sucrose: sucrose 1-fructosyltransferase (*Atq1SST-1*) cDNA amplification

Biological material

The *A. tequilana* FT sucrose: sucrose 1-fructosyltransferase (*Atq1SST-1*) isoform (JN790053, genbank accession code) was chosen as the target cDNA for heterologous expression in Arabidopsis. In order to amplify this FT from *A. tequilana*, several tissues including, leaves, stem and floral tissues were collected, immediately frozen using liquid nitrogen and stored at -70 °C to prevent biomolecule degradation (Figure 18).

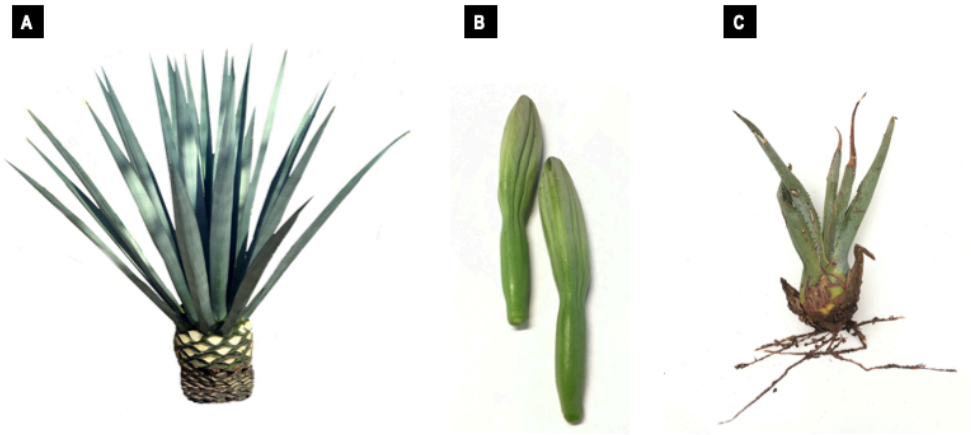


FIGURE 18. Biological tissues sampled from *A. tequilana* plants. **A**. Mature *A. tequilana* plant, **B**. Floral buds and **C**. *A. tequilana* bulbil.

RNA extraction

One hundred mg of each sample were ground mechanically by using liquid nitrogen and a mortar and pestle. As a first step, RNA extraction was carried out using 1 mL of Trizol reagent (Invitrogen; Carlsbad, CA, USA) homogenizing and leaving the samples at room

temperature for 5 min. Then, 200 µL of chloroform were added, the samples were homogenized, and left for 3 min at room temperature. The samples were then centrifuged at 12000 rpm for 15 min at 4 °C in order to separate the liquid phases; the aqueous phase was collected to purify the RNA using the PureLink RNA mini kit (Invitrogen). Finally, RNA concentration and purity were determined by the 260/ 280 nm ratio using a Nanodrop 8000 spectrophotometer (Thermo Scientific; Wilmington, DE, USA).

Reverse transcriptase – Polymerase Chain Reaction (RT-PCR)

The reverse transcriptase (RT) reaction was carried out using 500 ng of RNA within a 6 µL volume (83.33 ng/ µL). Then, 0.5 µL of the reverse oligonucleotide (5 µM) were added (Table 2) along with 2 µL of RT buffer 10 X, 1 µL of dNTPs (20 mM) and 0.5 µL of RT enzyme (RevertAid Reverse Transcriptase, 200 U/µL, Thermo Scientific; Waltham, MA, USA), in a final reaction volume of 10 µL. The reaction was carried out as follows using a Veriti thermocycler (Applied Biosystems; Foster City, CA, USA) (Figure 19):

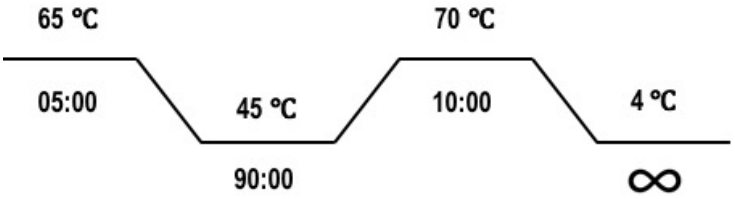


FIGURE 19. Schematic representation of the reverse transcriptase (RT) reaction thermal cycle.

In order to amplify the FT of interest, the PCR reaction was carried out by adding 2.5 µL of 10 X Taq Buffer with KCl (500 mM), 2 µL of $MgCl_2$ (25 mM), 1 µL of dNTPs (20 mM), 1 µL of forward oligonucleotide (10 µM), 1 µL of reverse oligonucleotide (10 µM) (Table 2), 0.25 µL of Taq polymerase (5U/ µL, Thermo Scientific), 3 µL of cDNA (dilution ratio 1: 10 v/ v) and MQ water in a final volume of 25 µL. The reactions were mixed and incubated

in a Veriti thermocycler (Applied Biosystems) following the thermal program shown below (Figure 20):

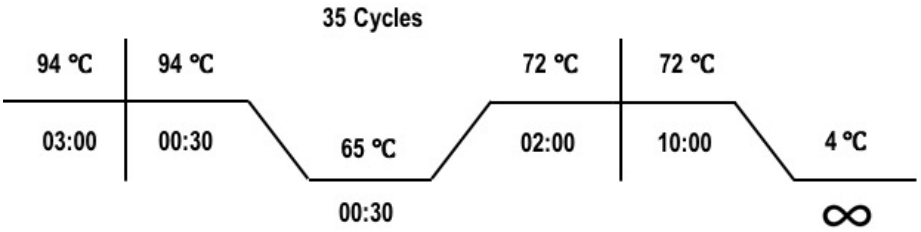


FIGURE 20. Schematic representation of the polymerase chain reaction (PCR) thermal cycle.

PCR reaction products were visualized by gel electrophoresis, loading 4 µL of PCR product plus 3 µL of loading buffer (Orange G; Sigma Aldrich, Steinheim, Germany) into 1% agarose gel dyed with 10000 X RedGel (Biotium; Fremont, CA, USA) and run at 70 V for 40 min. The PCR product purification was done by using ZymocleanTM Gel DNA Recovery Kit (Zymo Research; Irvine, CA, USA).

TABLE 2. List of oligonucleotides used to perform reverse transcriptase and polymerase chain reaction analysis.

Primer ID	Sequence 5' – 3'	Tm
FwattBSST1	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGCCTCATCGACAAAGGA	65 °C
RvattBSST1	GGGGACCACTTTGTACAAGAAAGCTGGGTCTCATGGAGCCACCAAGCCAG	
Actin 2 Fw	GTACAACCGGTATTGTGCTGGAT	65 °C
Actin 2 Rv	GCTTGGTGCAAGTGCTGTGATTTC	

Gateway cloning reaction and expression vector construction

Vector construction was carried out using the Gateway cloning technology (Invitrogen) based on the site-specific recombination properties of the lambda bacteriophage. Each cloning reaction was performed as described below.

BP cloning reaction

The BP cloning reaction requires an entry vector with the site-specific attP recombination sequences included and the amplified cDNA of interest with site-specific attB recombination sequences included within 5' and 3' ends. The reaction was performed as follows: 150 ng of purified cDNA, 1 µL of pDONR222 vector (150 ng/µL) (Invitrogen) (Figure 21), 2 µL of BP clonase II (Invitrogen) and water in a 10 µL volume reaction. The reaction was incubated at 25 °C for more than 16 h. The reaction was stopped by adding 1 µL of proteinase K and incubating at 37 °C for 10 min. Further manipulation of the reaction product was done on ice.

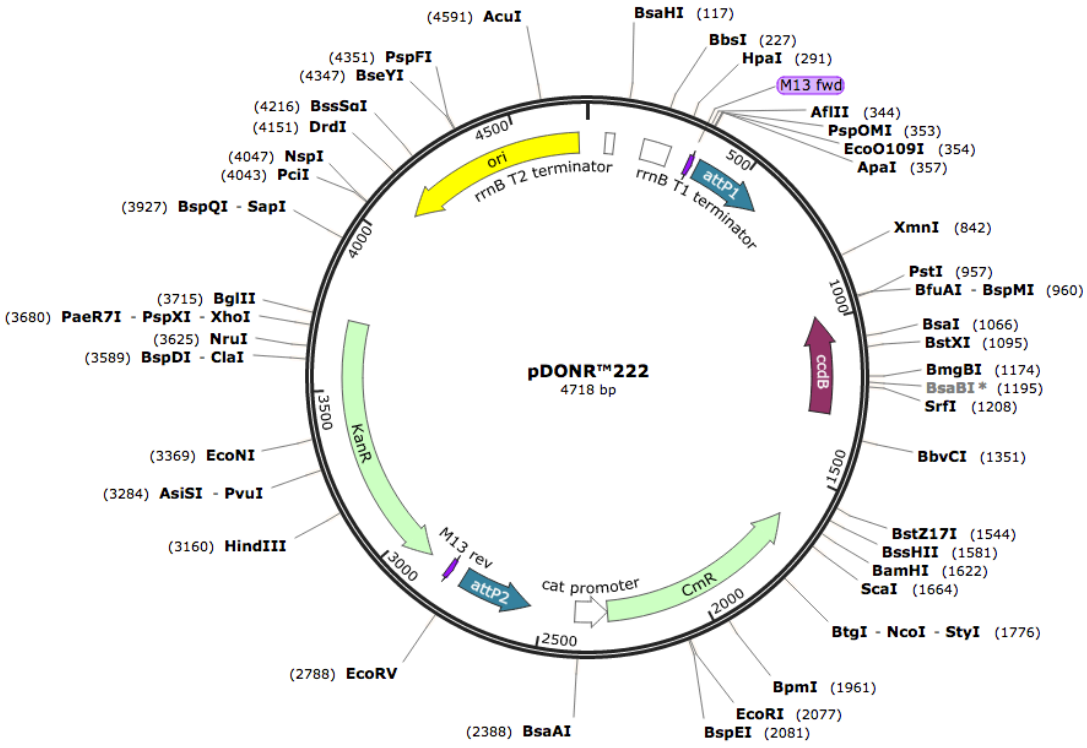


FIGURE 21. Schematic representation of the pDONR222 entry vector features. The vector is characterized by the presence of the recombination sites (attP1 AND attP2 sites) at the edges of a chloramphenicol resistance gene (*CmR*) and the *ccdB* lethal gene. Finally, the vector presents a kanamicyn resistance gene *(KanR*) and an origin of replication (ori). This vector is needed in order to generate the attL cloning recombination sites at the edges of the transgene for further recombination.

Escherichia coli (*E. coli*) transformation

E. coli DH10B competent cells (Invitrogen) were transformed through electroporation by using 2 µL of BP cloning reaction and adding 30 µL of sterile cold glycerol at 10% and 20 µL of electrocompetent cells into a 0.1 cm electroporation cuvettes (Invitrogen). After the voltage shock, the bacteria were re-suspended immediately in 500 µL of SOC media (Invitrogen) and incubated at 37 °C for 1 h. One hundred µL of re-suspended cells were plated in LB solid media with kanamycin (50 µM/ mL) and incubated at 37 °C for 16 h. Several colonies were chosen to confirm vector insertion into bacteria and the correct cDNA recombination into the pDONR222 vector by restriction enzyme assays. Positive colonies were sequenced to compare and verify the cDNA clone sequence.

LR cloning reaction

The LR cloning reaction requires the entry vector whose cDNA of interest was recombined, and the attL recombination sites were generated at the edges of the transgene, and the destination vector with the attR recombination sites included. The reaction was performed as follows: 150 ng of entry vector, 150 ng of destination vector, pB7WG2D expression vector (Karimi, *et al.,* 2002) (Figure 22), 2 µL of LR clonase I (Invitrogen) and water in a 10 µL volume reaction. The reaction was incubated at 25 °C for more than 16 h, and it was stopped by adding 1 µL of proteinase K incubating at 37 °C for 10 min. Further manipulation of the reaction product was done on ice.

Agrobacterium tumefaciens (*A. tumefaciens*) transformation

A. tumefaciens strain GV2260 was transformed by using 2 µL of plasmid DNA extracted from *E. coli* and adding 30 µL of sterile cold glycerol at 10% and 20 µL of electrocompetent *A. tumefaciens* cells (strain GV2260) in 0.1 cm electroporation cuvettes (Invitrogen). After the electric voltage shock, the bacteria were re-suspended immediately in 500 µL SOC media (Invitrogen) and incubated at 28 °C for 2 h. Re-suspended bacteria were diluted in

sterile MQ water in a 1: 200 v/ v ratio due to a high rate of bacterial growth. Aliquots (100 µL) of diluted bacteria were plated in YEB solid media with the set of selective antibiotics (i. e., spectinomycin [100 mg/mL], carbenicillin [100 mg/mL] and rifampicin [100 mg/mL]) and incubated at 28 °C for 48 h. For vector construction confirmation, several colonies were selected and grown in YEB liquid media in order to extract plasmid DNA. The plasmid DNA was used as template for PCR to confirm cDNA insertion into the expression vector by using specific oligonucleotides (Table 2).

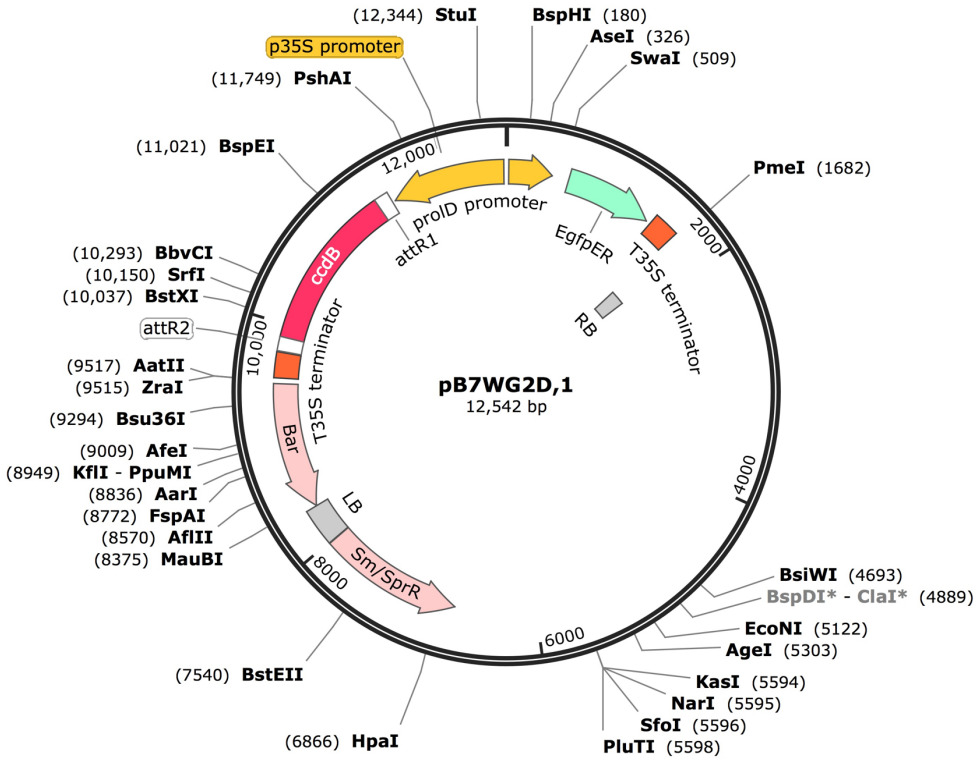


FIGURE 22. Schematic representation of the pB7WG2D expression vector features. The vector is characterized by the presence of spectinomicyn gene resistance (*Sm*) used for selection in prokaryote hosts. The T-DNA is characterized by the presence of green fluorescent protein marker (*EgfpER*), a cauliflower mosaic virus 35S promoter (P35S promoter), the *ccdB* lethal gene and glufosinate ammonium resistance gene (*bar*) used for selection in eukaryote hosts.

Arabidopsis Col. 0 transformation

Arabidopsis seed surface sterilization and growth

Arabidopsis ecotype Col-0 was selected as the wild type reference in order to carry out transformation and FT expression. Arabidopsis Col-0 seeds were superficially sterilized by washing them with a hypochlorite solution at 20%, then with an alcohol solution at 70% and finally three washes with deionized sterile water. Each wash was performed at 2000 rpm during 7 min in a ThermoMixer C (Eppendorf; Hamburg, Germany). After sterilization, the seeds were plated and spread in 0.1 X MS solid media and subjected to a vernalization period at 4 °C for 48 h, so all seeds could germinate uniformly. Finally, plates were placed in growth chambers at 22 °C under a 16 h light and 8 h dark photoperiod. Ten days after germination, the plantlets were transplanted into a soil mixture substrate for Arabidopsis composed of Sunshine, Perlite and Vermiculite (3: 1: 1). After approximately 40 days, the plants started to bloom, producing hundreds of floral buds ready to be transformed.

Floral dip

Before carrying out plant transformation through the floral dip method (Clough and Bent, 1998), *A. tumefaciens* cells were grown in YEB liquid media with the appropriate antibiotics (i. e., spectinomycin, carbenicillin and rifampicin) for 48 h. The cells were centrifuged at 12000 rpm for 3 min, discarding the supernatant and washing the cells three times with 0.5 X MS liquid media supplemented with sucrose at 5% (w/ v). Each wash was performed at 12000 rpm for 5 min. After the washes, the cells were re-suspended in 500 µL of 0.5 X MS liquid media supplemented with sucrose at 5% (w/v) and Silwet-L77 at 0.02% (v/v). The infiltration solution was inoculated to closed floral buds three days a week throughout the whole Arabidopsis floral cycle. For each infiltration day, new *A. tumefaciens* cells were grown and washed as mentioned above, using an inoculum from the previous cell culture.

Selection of transformed Arabidopsis lines

Due to the presence of the *bar* gene (resistance to glufosinate ammonium, a common chemical herbicide) within the T-DNA of the pB7WG2D expression vector, seeds harvested from agro-infiltrated Arabidopsis (T0) were superficially sterilized, spread and grown in 0.1 X MS solid media adding glufosinate ammonium (20 mg/ L) (Sigma-Aldrich; St. Louis, MO, USA). Plates were vernalized, as described previously, and were then placed in a growth chamber under a 16 h light and 8 h dark photoperiod. Ten days after the end of the vernalization period, the glufosinate resistance phenotype between transformed and non-transformed seeds could be detected. The resistant seedlings were transplanted into soil and monitored throughout their life cycle. The plants were grown in a growth chamber under the same photoperiod conditions. Leaf tissue from all the plants was sampled, frozen using liquid nitrogen and stored at -70°C until further analysis. Insertion of T-DNA into the plant genome was confirmed by PCR.

Analysis of gene expression and determination of water-soluble carbohydrate profiles in WT and transgenic Arabidopsis Col. 0 lines

Biological material

In order to determine gene expression and carbohydrate patterns in WT Arabidopsis Col-0, different tissues were collected throughout its life-cycle: i. e., plantlets (A), young leaves (B), vegetative leaves (C), senescent leaves (D), roots from vegetative plants (E), roots from reproductive plants (F), floral buds (G) and opened flowers (H) (Figure 23). All tissues were collected as a pool, frozen with liquid nitrogen and stored at -70 °C until analysis. All plants were grown in a growth chamber under a 16 h light and 8 h dark photoperiod and did not present any stress symptoms.

For transgenic lines, the leaves were the only tissue selected for both analyses, due to the presence of the cauliflower mosaic virus 35S (p35S) promoter, a constitutive promoter

that normally drives high levels of transgene expression. Plant growth conditions were as described previously. The samples were frozen using liquid nitrogen and stored at -70 °C until analysis.

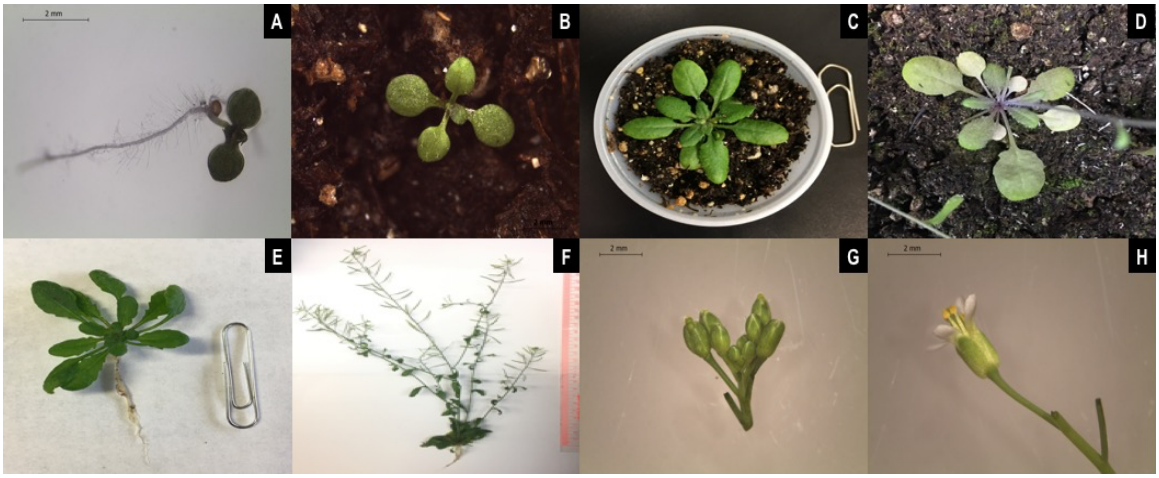


FIGURE 23. Biological tissues collected from WT Arabidopsis Col. 0 for qRT-PCR and carbohydrate profile analysis: **A**. Plantlet, **B**. Young leaves, **C**. Vegetative leaves, **D**. Senescent leaves, **E**. Roots from vegetative plants, **F**. Roots from reproductive plants, **G**. Floral buds and **H**. Opened flowers.

Gene expression analysis (qRT-PCR)

The oligonucleotide set used for quantitative analysis were designed using the primer design tool at NCBI (www.ncbi.nlm.nih.gov/tools/primer-blast/). The genes of interest (GOI) to be analyzed were cell wall invertase 3 (*Atcwinv3, 6-FEH*) (AT1G55120, Tair ID) and cell wall invertase 6 (*Atcwinv6, 6, 1-FEH*) (AT5G11920, Tair ID) from WT Arabidopsis Col. 0. Additionally, cell wall invertases 3 (*Atcwinv3*) and 6 (*Atcwinv6*) along with sucrose: sucrose 1-fructosyltranferase (*Atq1-SST1*) were analyzed in Arabidopsis transgenic lines. As a first step, the oligonucleotides set of GOI were tested for specificity by a real time PCR reaction. Then, a standard curve was performed using the oligonucleotides sets of GOI and *ubiquitin 11* (AT4G050501, Tair ID) as a housekeeping gene, in order to validate the GOI and their compatibility with the housekeeping gene (Table 3).

TABLE 3. List of oligonucleotide sequences used for qPCR analysis of the genes of interest.

Primer ID	Organism	Sequence 5' – 3'	Tm
FwCwinv3	*Arabidopsis thaliana*	TCTCAACCAACCGTACCGGA	65 °C
RvCwinv3		TTCTCACATCCCAAACGGCA	
FwCwinv6	*Arabidopsis thaliana*	TGGATGTGTTTACGGGCCTT	65 °C
RvCwinv6		TCTTCTCGCTGCTGCACATT	
FwSST1	*Arabidopsis thaliana*	GTTCCAGTGCTTCACGGTGA	65 °C
RvSST1		ACCCGTGCCGAGCTGTAGAT	
FwUbiquitin11	*Arabidopsis thaliana*	AGCGTCTCATCTTCGCTGGA	65 °C
RvUbiquitin11		AGAAACCACCACGGAGACGG	

Relative quantification experiments of the GOI were designed by using StepOne V2.3 software (Applied Biosystems) and gene expression analysis was done based on the comparative Ct method (ΔΔCt method) normalizing the data in function of *ubiquitin 11* expression. The reactions were performed as follows: 5 µL of SYBR Green PCR Master Mix (Applied Biosystems), 0.2 µL of forward oligonucleotide (10 µM), 0.2 µL of reverse oligonucleotide (10 µM), 1.1 µL of MQ H_2O and 3.5 µL of cDNA (synthesized from 500 ng of RNA from each sample) diluted in a 1: 5 v/ v ratio, having a volume reaction of 10 µL. Every reaction was done in duplicate with the respective "no template control" (NTC) for each gene. The final results showed a standard deviation from a technical triplicate. The reactions were carried out in a StepOne Plus thermocycler (Applied Biosystems) following the thermal oscillations shown below (Figure 24):

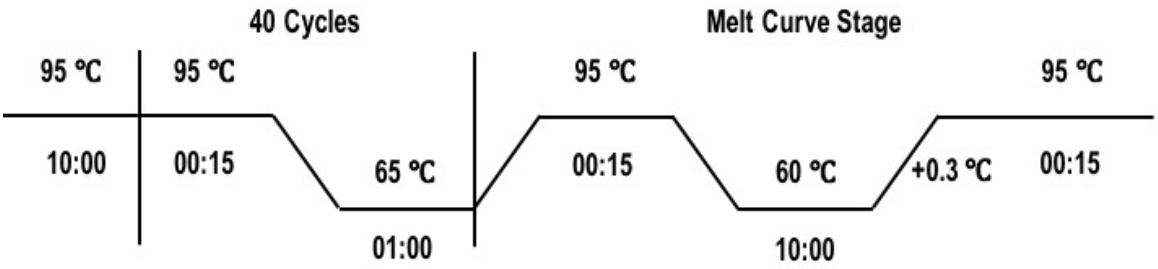


FIGURE 24. Schematic representation of the thermal conditions used to carry out qPCR and the melting curves.

Water-soluble carbohydrate extraction

Water-soluble extraction was done, according to Ávila de Dios and colleagues (2015) with some modifications, as follows: ~50 mg of ground tissue (fresh weight) were re-suspended in 1.5 mL of distilled water and incubated in a ThermoMixer C (Eppendorf) at 400 rpm, 70 °C for 30 min. Then, the samples were centrifuged at 14800 rpm for 10 min and the supernatant was recovered. The pellet was re-suspended in 1 mL of distilled water performing a second extraction incubating the samples again at 400 rpm, 70 °C for 15 min. The samples were centrifuged as before, recovering the supernatant, which was combined together with the previous one. Finally, the extracts were frozen with liquid nitrogen and placed into a freeze dryer (Labconco; Kansas City, MO, USA) for 24 h in order to dry the samples, producing a powder containing the soluble molecules present in the extracts.

Thin Layer Chromatography (TLC)

The extracted water-soluble carbohydrates were re-suspended in distilled water to a final concentration of 25 mg/ mL. One µL of each extract was placed on silica gel matrix plates (Sigma-Aldrich) together with a standard solution of a mixture of glucose, fructose, and sucrose (Sigma-Aldrich), DP3 (1-kestose), DP4 (nystose) and DP5 (fructofuranosyl-nystose) FOS (Wako Chemical; Chuo-Ku, Osaka, Japan), to identify the presence or

absence of fructan molecules in the samples. The plates were incubated and run three times (letting the plates dry between each run) in a chamber using as mobile phase a solvent solution of butanol- glacial acetic acid- distilled water (50: 25: 25 v/ v/ v). To visualize the carbohydrate profile of the samples, a post-chromatographic detection reaction was done using aniline- diphenylamine- phosphoric acid in acetone. Once the plate had dried, it was heated in an oven at 70 °C for 15 min to visualize the carbohydrate profile of the samples.

Results

Sucrose: sucrose 1-fructosyltransferase (*Atq1SST-1*) amplification and purification

The cDNA length *in silico* of *Atq1SST-1* has 1866 bp, but by adding the recombination attachment site sequences (attB) to the oligonucleotides (forward and reverse), the expected cDNA reached 1927 bp in length. Thus, RT-PCR was performed using RNAs from vegetative and floral tissues from *A. tequilana* plants. cDNA amplification was only visualized in the latter. The amplicon product had a size close to 2 Kb, assuming that the amplified product corresponded to *Atq1SST-1* (Figure 25, A). By using the same RNA, a second RT-PCR was carried out but in a higher reaction volume (50 µL) in order to increase the number of cDNA copies and once purified, to obtain a much higher cDNA concentration (Figure 25, B). Finally, the PCR product was purified by using and following the protocol by Zymo clean gel DNA recovery kit (Zymo Research). The purified *Atq1SST-1* cDNA was re-suspended in 6 µL of water, with a final concentration of 136.8 ng/ µL, which was enough to perform the subsequent cloning reactions.

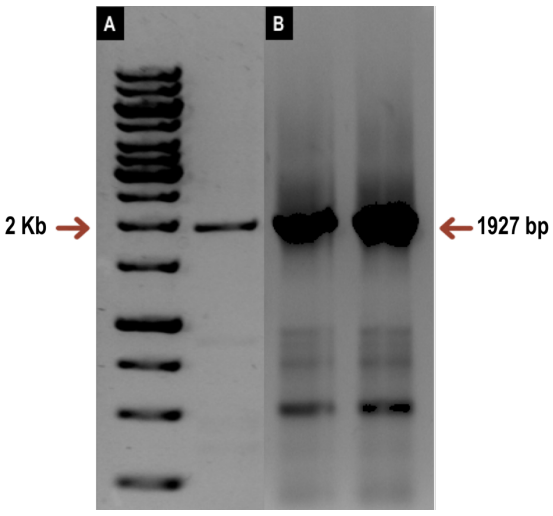


FIGURE 25. RT-PCR amplification analysis from *A. tequilana* tissues. **A.** *Atq1SST-1* cDNA amplification and **B**. *Atq1SST-1* cDNA amplification after loading 50 µl of the PCR reaction performed for subsequent cDNA purification.

BP cloning reaction

Once the *Atq1SST-1* cDNA was amplified with the corresponding attachment sites at the 5' and 3' edges, the first gateway cloning reaction was carried out along with the pDONR222 entry vector. The BP cloning reaction product was inserted into *E. coli* cells and grown in selection media (LB media) adding kanamycin due to the presence of the kanamycin resistance gene within the entry vector. Six colonies were selected randomly in order to confirm the correct cDNA recombination into the vector and the insertion of the constructed vector inside the cell host by restriction enzyme assays.

In silico simulation of enzyme restriction digestion was performed using SnapGene software in order to identify which enzymes cut the transgene sequence and the pDONR222 vector sequence. At the same time, a gel electrophoresis simulation of the selected restriction enzyme was done (Figure 26, A). This produced the corresponding bands pattern generated by a given endonuclease activity. The restriction enzymes selected were NcoI (10 U/μL, 500 U, Thermo Scientific) and EcoRV (10 U/μL, 2000 U, Thermo Scientific), able to cut inside and outside the transgene sequence, giving as a result bands of 2806 bp and 1564 bp, respectively.

Plasmid DNA extraction of the six colonies was performed in order to carry out the restriction enzyme assays. The assays were performed using 8 μL of plasmid DNA (less than 1 μg), 1 μL of buffer (the type of buffer depended on the restriction enzyme) and 1 μL of restriction enzyme. Each reaction was incubated for 3 h at 37 °C. Between each digestion, the reaction product was dialyzed to remove salts from the first buffer. As a result, a single colony (i. e., colony number 3) of the six analyzed was identified with the same band size pattern to the one predicted *in silico* (Figure 26, B) while the rest of the samples showed only one band after the digestion, meaning that the plasmid DNA had been linearized, only. As a negative control, the uncut, empty, vector was run as well.

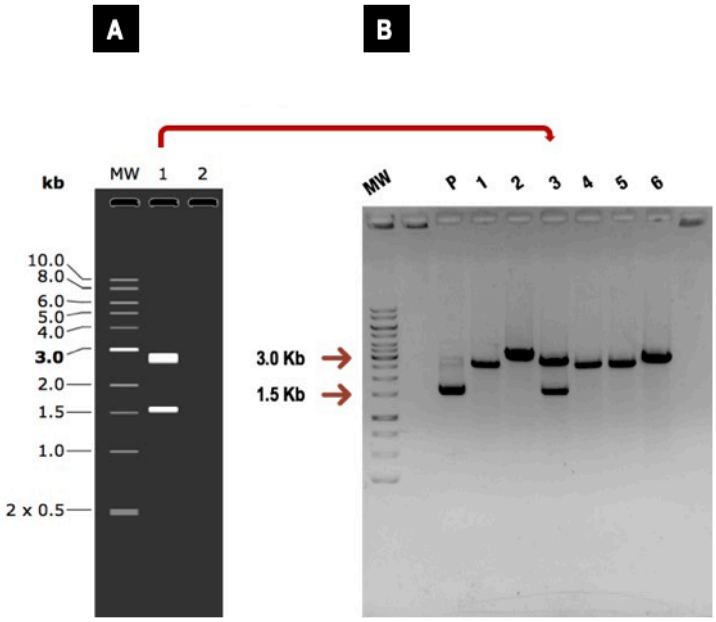


FIGURE 26. A. *In silico* gel electrophoresis simulation of the bands pattern of the pDONR222-*Atq1SST1* vector digested with the NcoI and EcoRV restriction enzymes. **B**. Gel electrophoresis of the products generated by NcoI and EcoRV endonuclease activity on the six *E. coli* colonies analyzed. MW. 1KB size ladder, P. Uncut pDONR222 VECTOR, 1-6. Code number of the colonies analyzed.

pDONR222 – *Atq1SST-1* sequencing

Having identified the colony whose plasmid DNA presented the same digestion pattern to the one simulated *in silico*, the next step was to confirm that the cDNA sequence from the amplified product corresponded to the one reported by Cortés-Romero and colleagues (2012) (JN790053, genbank accession code). Therefore, the plasmid DNA was sequenced using universal M13 oligonucleotides (forward and reverse) located at the ends of the transgene (Figure 20).

By aligning the nucleotide and amino acid sequences of the *Atq1SST-1* cDNA reported and the sequence derived from the transgene inserted into the pDONR222 vector, it was corroborated that both sequences were highly similar. However, the sequence obtained from the sequencing analysis was incomplete due to short sequence reads (about 800 nucleotides) and no internal primers were designed in order to get a full-length sequence.

At the nucleotide level, only 1577 nucleotides of the total 1866 nucleotides of the original cDNA could be sequenced. The sequences presented 98.5% of similarity, differing in only 26 nucleotides between them (Figure 27).

By translating the nucleotide sequences into amino acids, it was possible to obtain a partial protein sequence of 526 amino acids of the 621 amino acids which conform the Atq1SST-1 protein, presenting a 98.9% similarity between sequences. Although the sequence presented a stretch of almost 100 amino acids missing in the middle section, two of the three highly conserved motifs from PGHF32 reported by Ávila de Dios and colleagues (2015) in agave species were identified (Figure 28), ensuring that the amplified cDNA from *A. tequilana* most probably corresponds to FT *Atq1SST-1* isoform. The differences presented between the amplified and reported DNA sequences could be due to PCR or sequencing errors.

LR cloning reaction

After sequence analysis corroborated the right identity of the cDNA sequence of interest, the second cloning reaction was carried out. The LR cloning reaction was performed using the pDONR222-*Atq1SST-1* entry vector and the pB7WG2D expression vector. The product of the reaction was inserted into *E. coli* cells and grown in selective media (LB media), adding spectinomycin as the selective agent (considering that the pB7WG2D vector contains a spectinomycin resistance gene, Figure 22). The number of colonies obtained was low, since only 3 viable colonies were obtained.

```
SST1      ATGGCCTCATCGACAAAGGATGTGGAGGCTCCTCCAACTTTGGACGCCCCTCTCCTCGGG      60
Seq       ATGGCCTCATCGACAAAGGATGTGGAGGCTCCTCCAACTTTGGACGCCCCTCTCCTCGGG      60
          ************************************************************

SST1      TCTGCTGCCCCGCGGAGCAGGTTAAGGGTCGCCGCCGTCTCCCTCTCGGTCATGGCGTTT      120
Seq       TCTGCTGCCCCGCGGAGCAGGTTAAGGGTCGCCCCCGTCTCCCTCTCGGTCATGGCGTTT      120
          ********************************* ************************

SST1      CTTCTGGTGGCCATCGCGGCGGCGGTTCTCTACTACAACCCCGGCGGCGTCGCCTCAAAT      180
Seq       CTTCTGGTGGCCATCGCGGCGGCGGTTCTCTACTACAACCCCGGCGGCGTCGCCTCAAAT      180
          ************************************************************

SST1      CTGATGAGATTGAGAGAGAACGATTACCCGTGGACCAACGACATGCTGAGGTGGCAGCGC      240
Seq       CTGATGAGATTGAGAGAGAACGATTACCCGTGGACCAACGACATGCTGAGGTGGCAGCGC      240
          ************************************************************

SST1      ACTGGATTCCATTTCCAGCCAGAGAAAAATTTCCAGGCTGATCCCAATGCTGCCATGTTC      300
Seq       ACTGGATTCCATTTCCAGCCAGAGAAAAATTTCCAGGCTGATCCCAATGCTGCCATGTTC      300
          ************************************************************

SST1      TACAAGGGCTGGTATCATTTCTTCTACCAGTACAATCCGACAGGTGTAGCCTGGGACTAT      360
Seq       TACAAGGGCTGGTATCATTTCTTCTACCAGTACAATCCGACAGGCGTAGCCTGGGACTAT      360
          ******************************************* **************

SST1      ACAATTTCATGGGGCCACGCCGTTTCTAAGGATCTCCTCCATTGGAACTACCTCCCCATG      420
Seq       ACAATTTCATGGGGGCACGCGGTTTCTAAGGATCTCCTCCATTGGAACTACCTCCCCATG      420
          ************** ***** ************************************

SST1      GCTCTGAGGCCGGACCACTGGTACGACAGGAAAGGCGTCTGGTCGGGCTACTCCACGTTG      480
Seq       GCTCTGAGGCCGGACCACTGGTACGACAGGAAAGGCGTCTGGTCGGGCTACTCCACGTTG      480
          ************************************************************

SST1      TTGCCTGACGGTCGGATTGTGGTGTTGTACACCGGCGGTACCAAAGAATTAGTGCAAGTC      540
Seq       TTGCCTGACGGTCGGATTGTGGTGTTGTACACCGGCGGTACCAAAGAATTAGTGCAAGTC      540
          ************************************************************

SST1      CAAAACCTCGCGGTGCCCGTCAATCTCTCCGACCCTCTCCTCCTCGAGTGGAAGAAGTCT      600
Seq       CAAAACCTCGCCGTGCCCGTCAATCTCTCCGACCCTCTCCTCCTCGAGTGGAAGAAGTCT      600
          *********** ************************************************

SST1      CACGTCAATCCTATACTTGTTCCACCTCCTGGCATTGAAGACCACGACTTCAGGGATCCT      660
Seq       CACGTCAATCCTATACTTGTTCCACCACCTGGCATTGCAGTCAACGACTTCAGGGATCCT      660
          *************************** ********** ** * *****************

SST1      TTCCCAGTATGGTACAATGAGTCGGACTCCAGGTGGCATGTTGTGATTGGCTCCAAGGAT      720
Seq       TTCCCAGTCTGGTACAATGAGTCGGACTCCAGGTGGCATGTTGTGATTGGCTCCAAGGAT      720
          ******** ************************************************

SST1      CCGGAGCACTACGGCATTGTCCTCATCTACACCACCAAAGACTTCGTCAACTTCACTCTG      780
Seq       CCGGAGCACTACGGCATTGTCCTCATCTACACCACCAAAGATTTCGTCAACTTCACTCTG      780
          ************************************* ******************

SST1      CTCCCCAACATCCTCCACTCCACCAAGCAGCCCGTCGGCATGCTGGAATGCGTCGACTTG      840
Seq       CTCCCCAACATCCTCC----------------------------------------------      796
          ****************

SST1      TTCCCGGTCGCCACCACCGACTCCCGAGCCAACCAGGCCCTCGACATGACGACGATGAGG      900
Seq       ------------------------------------------------------------      796

SST1      CCGGGGCCGGGGTTGAAGTACGTGCTTAAAGCGAGCATGGACGACGAGAGGCACGATTAC      960
Seq       ------------------------------------------------------------      796

SST1      TATGCCCTCGGATCATTCGATCTGGACTCTTTCACTTTTACTCCCGACGATGAAACCATT     1020
Seq       ------------------------------------------------------------      796

SST1      GACGTTGGGATTGGATTGAGGTACGACTGGGGTAAATTTTATGCTTCCAAGACTTTCTAC     1080
Seq       ------------------------------------------------------------      796

SST1      GATCAGGAGAAGCAGAGGCGGGTGCTGTGGGGATACGTTGGGGAGGTCGATAGTAAGAGA     1140
Seq       -----GGAGAAGCAGAGGCGGGTGCTGTGGGGATACGTTGGGGAGGTCGATAGTAAGAGA      851
               ******************************************************

SST1      GATGACGCGTTGAAAGGTTGGGCCTCACTTCAGAACATTCCGCGAACAATATTGTTCGAT     1200
Seq       GATGACGCGTTGAAAGGTTGGGCCTCACTTCAGAACATTCCGCGAACGATATTGTTCGAT      911
          ********************************************* ************

SST1      ACGAAAACTAAAAGCAACCTCATCCTGTGGCCGGTTGAGGAAGTGGAGAGCCTCAGGACG     1260
Seq       ACGAAAACGAAAAGCAACCTCATCCTGTGGCCGGTTGAGGAAGTGGAGAGCCTCAGGACG      971
          ******** ************************************************
```

```
SST1      ATCAACAAAAATTTCAACAGCATCCCCCTGTATCCTGGGTCGACATACCAGCTAGATGTT    1320
Seq       ATCAACAAAAATTTCAACAGCATTCCCCTGTATCCTGGGTCGACATACCAGCTAGATGTT    1031
          ********************** ************************************

SST1      GGCGAAGCAACTCAGTTGGACATAGTAGCCGAGTTTGAGGTCGATGAGAAGGCGATAGAG    1380
Seq       GGCGAAGCAACTCAGTTGGACATAGTAGCCGAGTTTGAGGTCGATGAGAAGGCGATAGAG    1091
          ************************************************************

SST1      GCTACAGCCGAGGCCGACGTCACCTACAATTGCAGCACCAGCGGCGGTGCTGCCAACCGG    1440
Seq       GCTACAGCCGAGGCCGACGTCACCTACAACTGCAGCACCAGCGGCGGGGCAGCCAACCGG    1151
          **************************** ***************** ** *********

SST1      GGCGTTCTCGGGCCGTTCGGCCTGCTCGTTCTCGCGAATCAGGAACTCTCCGAGCAGACT    1500
Seq       GGCGTTCTCGGACCGTTCGGCCTGCTCGTTCTCGCGAATCAGGAACTCTCCGAGCAGACT     1211
          *********** ************************************************

SST1      GCGACCTACTTCTATGTTAGTCGGGGAATTGATGGCAATCTGCGCACTCACTTCTGCCAG    1560
Seq       GCGACCTACTTCTATGTTAGTCGGGGAATTGATGGCAATCTGCGCACTCACTTCTGCCAG    1271
          ************************************************************

SST1      GATGAACTGAGGTCATCCAAGGCTGGGGCCATCACTAAGAGAGTGGTTGGCAGCACCGTT    1620
Seq       GATGAACTGAGGTCACCCAAGGCTGGGGCCATCACCAAGAGAGTGGTTGGCAGCACCGTT    1331
          *************** ****************** ************************

SST1      CCAGTGCTTCACGGTGAAACCTGGGCATTAAGAATACTTGTGGATCACTCTATCGTGGAG    1680
Seq       CCAGTGCTTCACGGTGAAACCTGGGCATTAAGAATACTTGTGGATCACTCTATCGTCGAG    1391
          ********************************************************* ***

SST1      AGCTTTGCCCAGAGAGGGAGGGCGGTGGCCACGTCCCGTGTGTACCCGACTGAGGCAATC    1740
Seq       AGCTTTGCCCAGAGAGGGAGGGCGGTGGCCACATCCCGTGTGTACCCGACTGAGGCAATC    1451
          ******************************** ****************************

SST1      TACAGCTCGGCACGGGTCTTCCTCTTCAACAATGCAACCGACGCCATTGTCACCGCAAAA     1800
Seq       TACAGCTCGGCACGGGTCTTCCTCTTCAACAATGCAACCGACGCCATTGTCACCGCAAAA     1511
          ************************************************************

SST1      ACGGTGAACGTGTGGCACATAAACTCCACATACAACCATGTCTTCCCTGGCTTGGTGGCT    1860
Seq       ACGGTGAACGTGTGGCACATGAACTCCACATACAACCATGTCTTCCCTGGCTTGGTGGCT    1571
          ******************** ***************************************

SST1      CCATGA 1866
Seq       CCATGA 1577
          ******
```

FIGURE 27. *Atq1SST-1* nucleotide sequence alignment reported by Cortés-Romero and colleagues (2012) (SST1) and nucleotide sequence obtained from sequencing analysis (Seq). Asterisks and dashes indicate identical residues and missing residues, respectively.

```
SST1      MASSTKDVEAPPTLDAPLLGSAAPRSRLRVAAVSLSVMAFLLVAIAAAVLYYNPGGVASN      60
Seq       MASSTKDVEAPPTLDAPLLGSAAPRSRLRVAPVSLSVMAFLLVAIAAAVLYYNPGGVASN      60
          ****************************** ****************************

SST1      LMRLRENDYPWTNDMLRWQRTGFHFQPEKNFQADPNAAMFYKGWYHFFYQYNPTGVAWDY      120
Seq       LMRLRENDYPWTNDMLRWQRTGFHFQPEKNFQADPNAAMFYKGWYHFFYQYNPTGVAWDY      120
          ************************************************************

SST1      TISWGHAVSKDLLHWNYLPMALRPDHWYDRKGVWSGYSTLLPDGRIVVLYTGGTKELVQV      180
Seq       TISWGHAVSKDLLHWNYLPMALRPDHWYDRKGVWSGYSTLLPDGRIVVLYTGGTKELVQV      180
          ************************************************************

SST1      QNLAVPVNLSDPLLLEWKKSHVNPILVPPPGIEDHDFRDPFPVWYNESDSRWHVVIGSKD      240
Seq       QNLAVPVNLSDPLLLEWKKSHVNPILVPPPGIAVNDFRDPFPVWYNESDSRWHVVIGSKD      240
          ******************************  :************************

SST1      PEHYGIVLIYTTKDFVNFTLLPNILHSTKQPVGMLECVDLFPVATTDSRANQALDMTTMR      300
Seq       PEHYGIVLIYTTKDFVNFTLLPNILX----------------------------------      266
          *************************

SST1      PGPGLKYVLKASMDDERHDYYALGSFDLDSFTFTPDDETIDVGIGLRYDWGKFYASKTFY      360
Seq       ------------------------------------------------------------      266

SST1      DQEKQRRVLWGYVGEVDSKRDDALKGWASLQNIPRTILFDTKTKSNLILWPVEEVESLRT      420
Seq       -XEKQRRVLWGYVGEVDSKRDDALKGWASLQNIPRTILFDTKTKSNLILWPVEEVESLRT      325
            ************************************************************

SST1      INKNFNSIPLYPGSTYQLDVGEATQLDIVAEFEVDEKAIEATAEADVTYNCSTSGGAANR      480
Seq       INKNFNSIPLYPGSTYQLDVGEATQLDIVAEFEVDEKAIEATAEADVTYNCSTSGGAANR      385
          ************************************************************

SST1      GVLGPFGLLVLANQELSEQTATYFYVSRGIDGNLRTHFCQDELRSSKAGAITKRVVGSTV      540
Seq       GVLGPFGLLVLANQELSEQTATYFYVSRGIDGNLRTHFCQDELRSPKAGAITKRVVGSTV      445
          ****************************************** **************

SST1      PVLHGETWALRILVDHSIVESFAQRGRAVATSRVYPTEAIYSSARVFLFNNATDAIVTAK      600
Seq       PVLHGETWALRILVDHSIVESFAQRGRAVATSRVYPTEAIYSSARVFLFNNATDAIVTAK      505
          ************************************************************

SST1      TVNVWHINSTYNHVFPGLVAP   621
Seq       TVNVWHMNSTYNHVFPGLVAP   526
          ******:***************
```

FIGURE 28. Atq1SST-1 amino acid sequence alignment reported by Cortés-Romero and colleagues (2012) (SST1) compared to that obtained from the translation of the nucleotide sequence of the sequenced construct (Seq). Two of three PGHF32 conserved motifs are marked with red. The transgene shows the same motifs as those reported by Ávila de Dios and colleagues (2015). Asterisks, colons and dashes indicate identical residues, conserved substitutions, and missing residues, respectively.

As in the previous cloning reaction, *in silico* simulation of restriction enzyme digestion and gel electrophoresis were implemented, obtaining a bands pattern that could identify correct transgene recombination into the expression vector (Figure 29, A). The restriction enzymes selected were XbaI (10 U/ µL, Invitrogen) with a double cut and XhoI (4000 U, Invitrogen) with a unique cut, producing a 3-bands pattern of 1567 bp, 3094 bp and 8095 bp for the constructed expression vector. As negative control, empty pB7WG2D vector lacking the XhoI restriction site was digested using the same enzymes, generating a 2-bands pattern of 4447 bp and 8095 bp instead of the 3-bands pattern. The restriction enzyme reaction was performed using 8 µL of plasmid DNA (less than 1 µg of DNA), 1 µL of buffer and 0.5 µL of each enzyme.

As a result, the empty vector showed the 2-bands predicted, with the expected sizes resulting from the double cut catalyzed by the XbaI endonuclease enzyme (Figure 29, B). On the other hand, plasmid DNA from *E. coli* colonies, showed the expected bands pattern predicted by the *in silico* simulation of the transgene recombination, producing the expected 3-bands having the right sizes (Figure 29, B). This result confirmed the insertion of *Atq1SST-1* cDNA into the expression vector and also into *E. coli* cells.

A second alternative for transgene recombination corroboration was also carried out by PCR amplification. In this experiment, the *Atq1SST-1* transgene was amplified using plasmid DNA from *E. coli* colonies and as a negative control the pB7WG2D empty vector was used. A unique amplified band from each *E. coli* colony analyzed was produced. Their sizes were close to 2 Kb, while the no amplification was produced using the empty vector as template. This corroborated the proper transgene recombination into the expression vector and the correct expression vector construction (Figure 29, C).

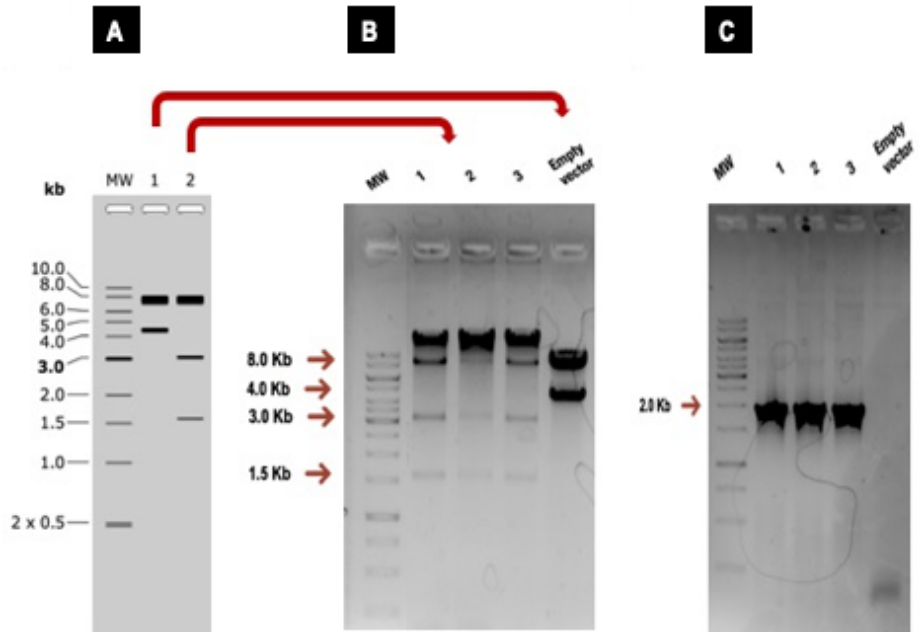


FIGURE 29. A. *In silico* gel electrophoresis of pB7WG2D **(1)** and pB7WG2D-*Atq1SST-1* **(2)** vector with the corresponding bands generated by XbaI and XhoI endonuclease activity. **B.** Gel electrophoresis separation of the products generated by the reaction of the same restriction enzymes on the plasmid DNA obtained from the three *E. coli* colonies analyzed. **C.** PCR product of *Atq1SST-1* using as template plasmid DNA from the three *E. coli* colonies analyzed and plasmid DNA from the empty vector. MW. 1KB size ladder, 1-3. Code number of the colonies analyzed.

Finally, the pB7WG2D-*Atq1SST-1* expression vector was inserted into *A. tumefaciens* strain GV2260. To accomplish this, *A. tumefaciens* cells were transformed using pB7WG2D-*Atq1SST-1* plasmid DNA and grown in selective media (YEB media) adding spectinomycin (due to the presence of resistance gene within the expression vector), carbenicillin and rifampicin (the natural antibiotic resistance of the GV2260 strain). Nine cell colonies were selected randomly to analyze expression vector insertion into the host. By PCR amplification of *Atq1SST-1* using as positive control the constructed vector isolated from *E. coli* cells, it was possible to confirm expression vector insertion in all nine *A. tumefaciens* colonies analyzed by PCR, which amplified a PCR product of about 2 Kb (Figure 30). Thus, one of this group of colonies was chosen for Arabidopsis transformation.

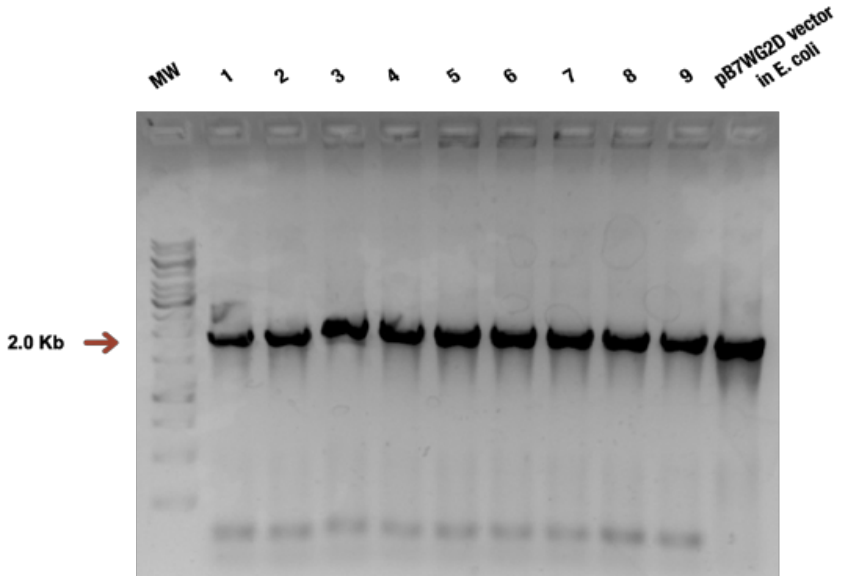


FIGURE 30. *Atq1SST-1* cDNA amplification of nine *A. tumefaciens* colonies analyzed. The pB7WG2D- *Atq1SST-1* construction isolated from *E. coli* cells was used as a positive control. MW. 1KB size ladder, 1-9. Code number of the colonies analyzed.

Arabidopsis transformation

Arabidopsis Col. 0 transformation was carried out according to the floral dip method, which required plants in blossom in order to infect the floral buds with *A. tumefaciens* cells carrying the pB7WG2D-*Atq1SST-1* expression vector. The plants were transformed throughout the Arabidopsis flowering stage (T0 generation) and when the Arabidopsis life cycle ended, the seeds were harvested. These represented the T1 generation. Approximately 2400 seeds were grown in MS media supplemented with glufosinate ammonium as the selection reagent (due to the presence of herbicide resistance gene, *bar*, within the T-DNA) in order to identify and select the seeds that were successfully transformed by showing an herbicide resistance phenotype. Ten days after the vernalization period, most of the non-transformed plants started to die because of the herbicide and only just a few plantlets survived showing an herbicide resistance phenotype, meaning that those plants were, most likely, the transformed ones.

The plantlets with the herbicide resistance phenotype were transferred to a new plate with solid MS media added with glufosinate ammonium in order to maintain selection and confirm herbicide resistance. In total, 42 plantlets transferred showed the capacity to grow in presence of the herbicide on the selective media (Figure 31).

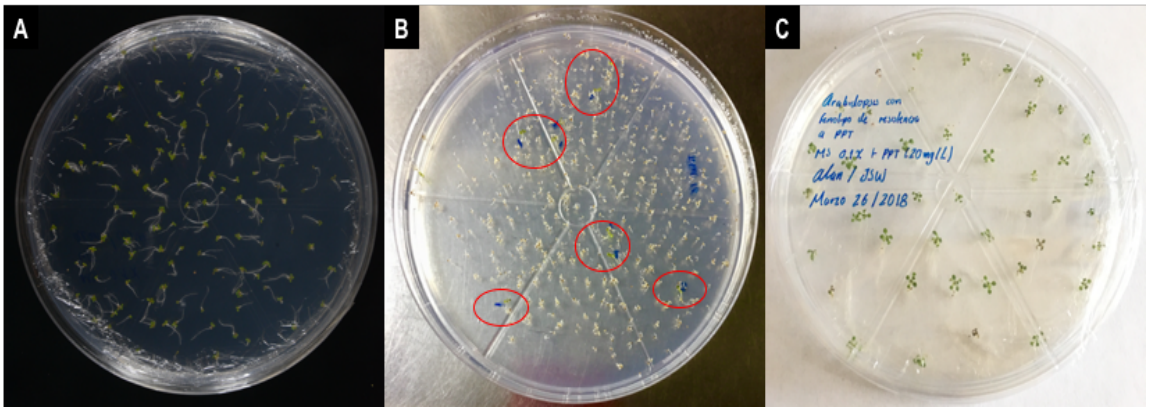


FIGURE 31. Arabidopsis Col. 0 seed germination. **A.** WT Arabidopsis seeds germinating on MS media, **B**. Arabidopsis seeds harvested from floral dipped plants germinating on MS media with glufosinate ammonium (20 mg/L) and **C.** Arabidopsis plantlets that showed an herbicide resistance phenotype (marked with red ovals in figure B), transplanted into MS media with glufosinate ammonium (20 mg/L).

When the resistance phenotype plantlets developed the first true leaves, they were transplanted into soil and kept under a high humidity environment. When the number of leaves increased in the rosette, all plants were sampled in order to confirm T-DNA insertion into the plant genome. Genomic DNA extraction was carried out in order to perform a PCR amplification of the *Atq1SST-1* transgene, using as negative control genomic DNA from WT Arabidopsis and as a positive control the pB7WG2D-*Atq1SST-1* plasmid DNA from *A. tumefaciens* cells.

From the 42 herbicide resistance plants, only 35 of them amplified a single product of almost 2 Kb, the same as the positive control while WT Arabidopsis DNA did not amplify any product, meaning that the T-DNA within *Atq1SST-1* transgene was inserted into the plant genome of those plants while 5 plants were distinguished as false positive plants

(Figure 32). Lastly, two plants died during the transition of transplanting to soil and it was therefore impossible to determine if they were transformed or not. The transformation efficiency achieved was of 1.4 %, a result that is within the efficiency range for the floral dip method (0.05 - 3.0%; Clough and Bent, 1998). The plants were kept under the same growth conditions and watered until seeds were ready to harvest (T2 generation).

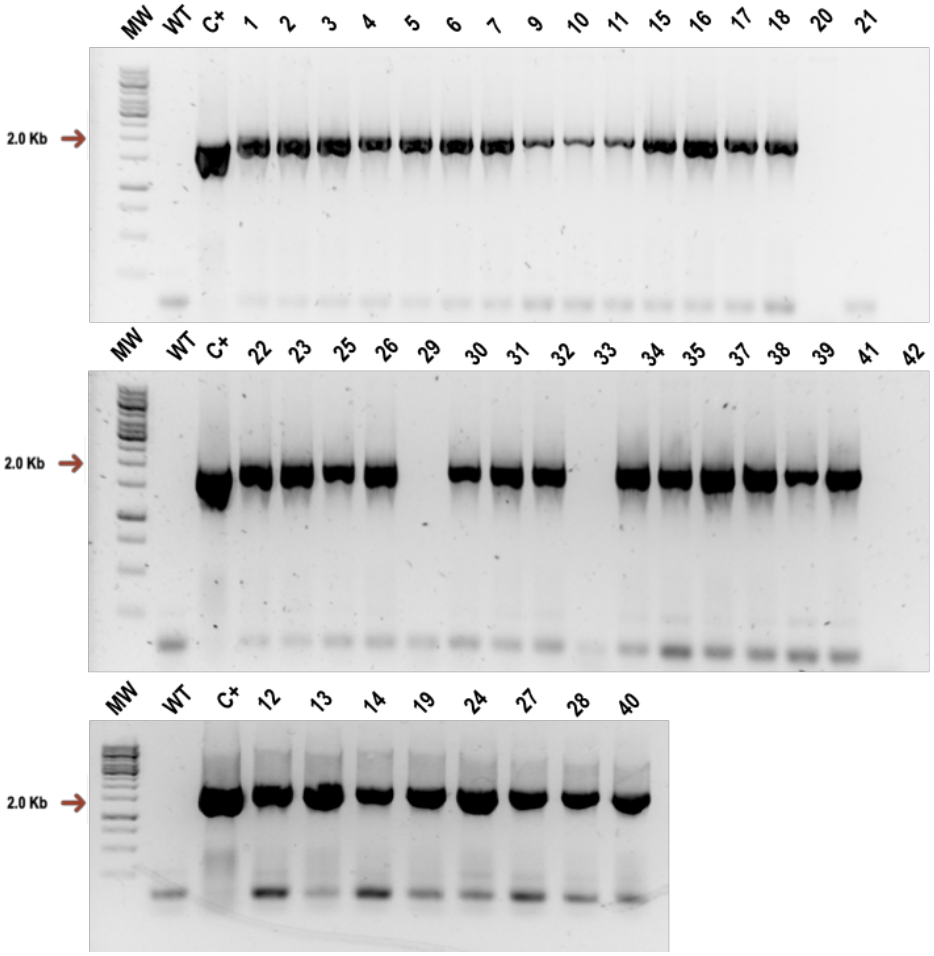


FIGURE 32. Gel electrophoresis of the PCR products corresponding to the *Atq1SST-1* transgene in Arabidopsis plants showing a glufosinate ammonium resistance phenotype. As negative control, genomic DNA of wild type Arabidopsis (WT) was used as template and as positive control plasmid DNA of *A. tumefaciens* carrying the expression vector was used as template (C+). MW. 1KB size ladder. 1-42. Code number of the plants analyzed.

qRT-PCR analysis of transgenic Arabidopsis lines harbouring *Atq1SST-1* transgene

In order to confirm functional transgene transcripts in transgenic lines, gene expression of *Atq1SST-1* was carried out and analyzed in 10 transformed Arabidopsis lines, chosen randomly from the 35 originally identified. Gene expression was quantified by extraction of RNA from leaves at the early reproductive stage from transformed Arabidopsis. The *Atq1SST-1* transgene is driven by the constitutive promoter cauliflower mosaic virus 35S (*p35S-Atq1SST-1*) and gene expression should be constitutive over all tissues, therefore only leaf tissue was analyzed.

As expected, the 10 transformed lines showed differences in gene expression levels for the *p35S-Atq1SST-1* transgene. The highest level of expression was observed in Arabidopsis line 5 followed by lines 6, 35 and 1 respectively, while all other lines showed approximately the same gene expression pattern which coincided with the lowest levels of expression identified in transgenic lines (Arabidopsis line 9 was the reference RNA, which means that it was the line with the lowest transgene expression of the 10 lines analyzed) (Figure 33).

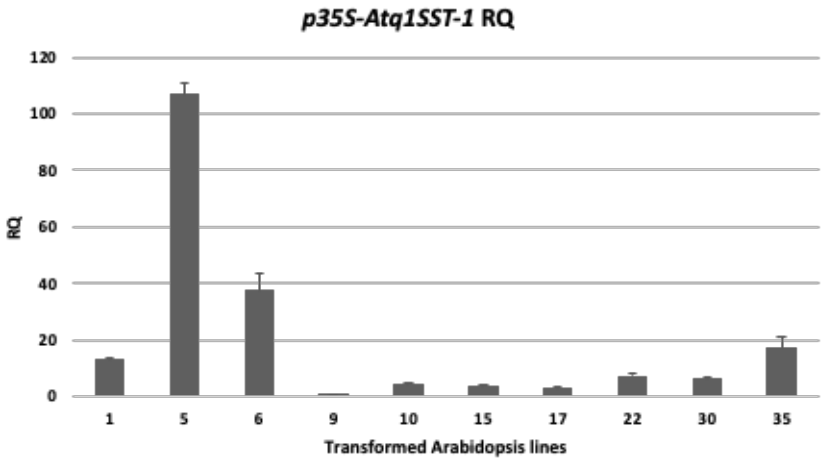


FIGURE 33. Relative gene expression quantification of *p35S-Atq1SSS-1* in different transgenic Arabidopsis lines. **RQ**. Relative quantification, and the numbers represent the different transgenic Arabidopsis lines that were analyzed.

Phenotype of transgenic Arabidopsis lines

The transformed lines showed no aberrant phenotype at the morphological level. Thus, the leaves, rosette and inflorescences presented a normal morphology, similar to WT Arabidopsis. Nevertheless, by comparing the transformed Arabidopsis lines to a WT Arabidopsis reference, the transformed plants presented an irregular development during flowering with shorter inflorescences and fewer branches in the inflorescence (Figure 34). This suggests that *p35S-Atq1SST-1* overexpression and putative deregulation of the endogenous *FEH* genes leads to these phenotypic changes in inflorescence development.

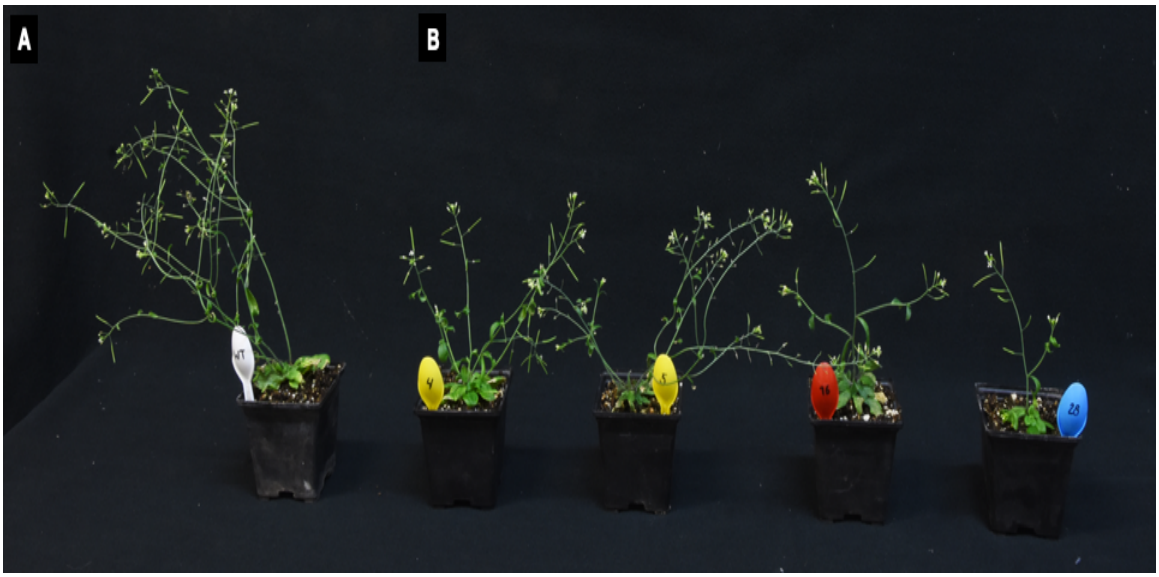


FIGURE 34. Phenotype comparison of **A.** WT Arabidopsis and **B**. A sample of the population of transformed lines generated.

Water-Soluble carbohydrate profile in WT and transgenic Arabidopsis Col.0

To determine whether expression of *p35S-Atq1SST-1* in transgenic Arabidopsis lines could produce an active 1-SST enzyme and lead to the synthesis of 1-kestose and also to confirm

the absence of fructan molecules in WT Arabidopsis, water soluble carbohydrates were extracted from several tissues of WT Arabidopsis plants throughout their life cycle, including plantlets, leaves (young, mature and senescent), roots (from vegetative and reproductive stages), floral buds and fully opened flowers, while in transgenic Arabidopsis lines the leaves were used. The presence and type of carbohydrate in the extracts was determined by TLC. The water-soluble carbohydrate profile was similar in most tissues for WT samples, showing glucose, fructose and sucrose but no evidence for the presence of 1-kestose (the simplest inulin fructan type molecule) (Figure 35). In some tissues such as mature leaves, senescent leaves, floral buds and fully opened flowers, a carbohydrate spot close to the DP4 FOS standard was identified. However, this spot presented a bluish color, indicating that the molecules most probably correspond to an oligosaccharide having a glucose backbone. Floral buds and fully opened flowers were the only tissues that showed additional oligosaccharide spots with a higher DP. Nevertheless, it was not possible to determine the nature of each of the unknown carbohydrates due to lack of adequate carbohydrate standards. However, the results strongly suggest that Arabidopsis does not synthesize/ accumulate fructans.

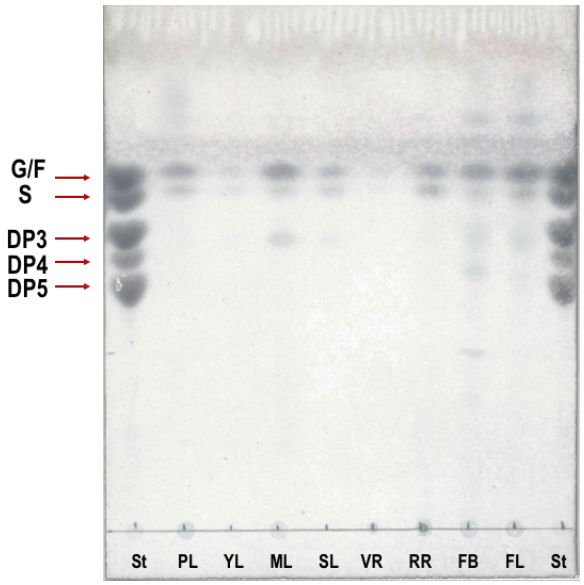


FIGURE 35. Water-soluble carbohydrate profile determined by TLC from WT Arabidopsis tissues. **St**. Standard: glucose (G), fructose (F), sucrose (S), 1-kestose (DP3), nystose (DP4) and fructofuranosyl-nystose (DP5), **PL**. Plantlet, **YL**. Young leaf, **ML**. Mature leaf, **SL.** Senescent leaf, **VR**. Vegetative root, **RR.** Reproductive root, **FB**. Floral bud and **FL**. Fully developed flower.

Water soluble carbohydrates were also extracted and analyzed by TLC from leaves of the 10 transformed Arabidopsis selected for qPCR analysis. None of the transformed plants showed evidence of fructan synthesis (Figure 36). In both transgenic and WT plants, the only carbohydrate spots identified were glucose, fructose and sucrose. Additionally, a bluish spot (probably maltose) was identified below sucrose spot. A possible explanation is that 1-kestose accumulates at such low levels in the transgenic lines that it is below the level of detection by TLC.

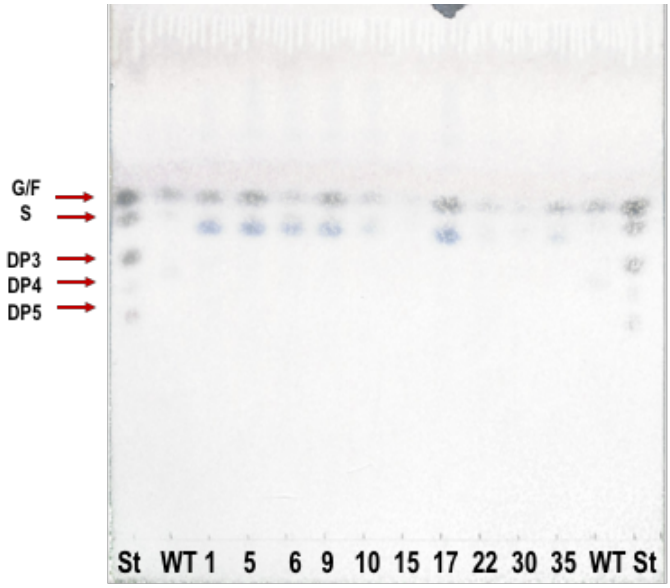


FIGURE 36. Water soluble carbohydrate profile determined by TLC analysis of leaf extracts obtained from transformed Arabidopsis plants. **St**. Standards: glucose (G), fructose (F), sucrose (S) 1-kestose (DP3), nystose (DP4) and fructofuranosyl-nystose (DP5), **WT**. WT Arabidopsis; the numbers represent each of the transformed Arabidopsis lines analyzed.

In order to address this possibility, extracts were obtained from whole plants and analyzed by TLC. Only two plants, which were still in the reproductive stage and most of the rosette leaves were green, were selected for this experiment and as control, a whole WT Arabidopsis plant was analyzed to compare the carbohydrate profiles (Figure 37, A).

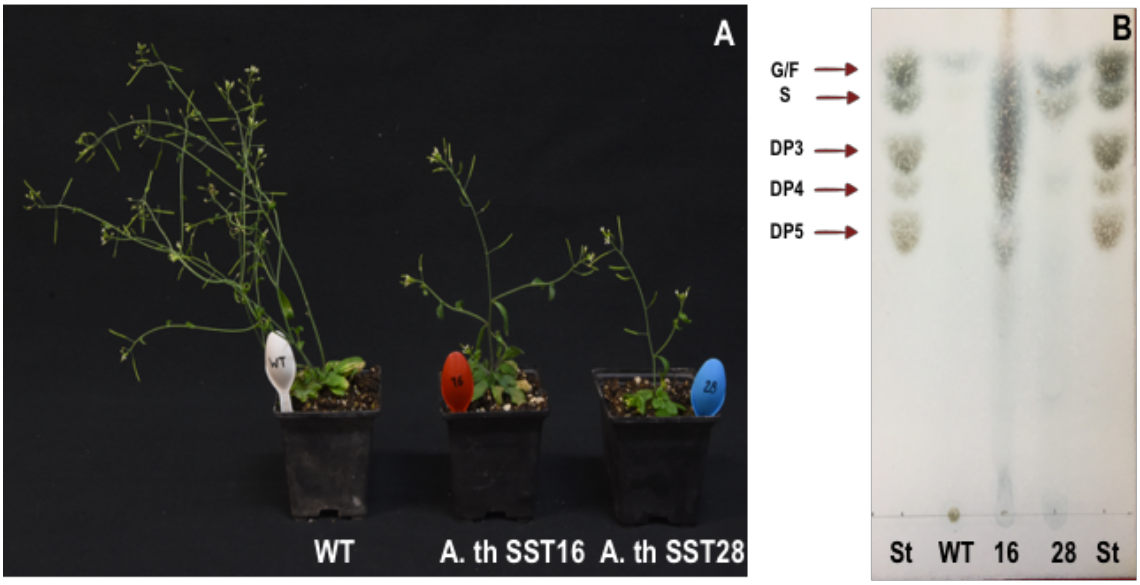


FIGURE 37. A. WT Arabidopsis reference and transformed lines selected for whole plant extraction analysis, **B**. TLC analysis of wild type Arabidopsis (WT) and whole plant extracts corresponding to transgenic lines 16 and 28. St. Standard: glucose (G), fructose (F), sucrose (S) 1-kestose (DP3), nystose (DP4) and fructofuranosyl-nystose (DP5).

TLC analysis of whole plant extract did not show a convincing 1-kestose spot below the sucrose spot as expected (Figure 37, B). On the contrary, sample 16 only displayed a large spot with no separation of monosaccharides and disaccharides while sample 28 did show an appreciable separation of the simpler carbohydrates. Nevertheless, a spot was identified near DP5 in sample 16, and near DP4 in sample 28, that might correspond to some FOS. In order to confirm the TLC result, the extracts were also analyzed by High-performance anion-exchange chromatography with pulsed amperometric detection (HPAEC-PAD) in the laboratory of Dr. Win Van den Ende in Louvain, Belgium. As can be observed, neither 1-kestose nor nystose (DP4) were determined to be present in any of the transgenic lines (Figure 38) tested. Although, in contrast to the WT line, a strong signal for maltose was detected.

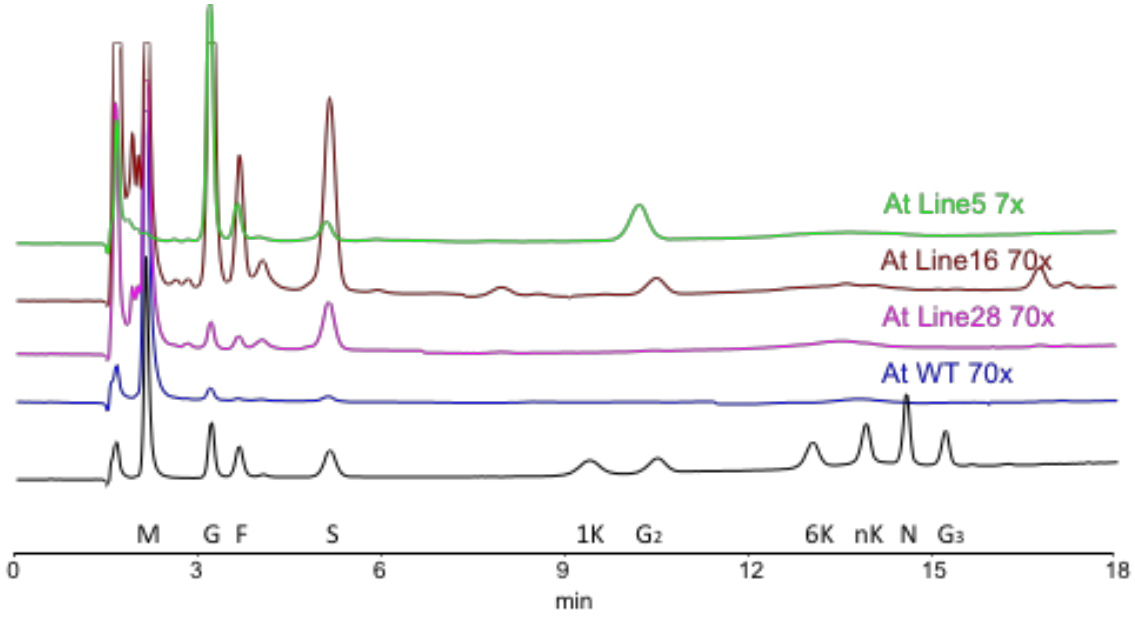


FIGURE 38. High-performance anion-exchange chromatography with pulsed amperometric detection (HPAE-PAD) of WT and transgenic Arabidopsis. M. Mannitol (internal standard), G. Glucose, F. Fructose, S. Sucrose, 1K. 1-Kestose, G2. Maltose, 6K. 6-Kestose, NK. Neokestose, N. Nystose and G3. Maltotriose. WT. Wild type extract and the "5", "16" and "28" numbers represent the transgenic lines extracts that were analyzed.

In silico gene expression analysis of cell wall invertase 3 (*6-FEH*) and cell wall invertase 6 (*6, 1-FEH*) in WT Arabidopsis

Due to lack of expression analysis reports for the two endogenous *FEH* genes in Arabidopsis, the first task was to consult a data base of *in silico* gene expression for both *FEH* in order to obtain an idea of the expression patterns in different tissues and developmental stages of Arabidopsis (Schmid et al., 2005). The use of this information had the purpose of selecting the best tissues and the development stages for further analysis in the transformed plants. The *in silico* gene expression pattern was consulted at the ePlant database (bar.utoronto.ca, Winter, *et al.,* 2007) and expression patterns of the two *FEH* were obtained indicating that their expression levels were completely different.

According to the database, cell wall invertase 3 (*Atcwinv3*, *6-FEH*, AT1G55120) is more highly expressed in comparison to the *Atcwinv6*. The tissues where the highest expression

of *Atcwinv3* was found included mature floral buds, floral tissues in specific petals and stamens followed by sepals and carpels and lastly, in the first stages of embryo development. The tissues where *Atcwinv3* was moderately expressed were shoot apical meristems as well as leaves throughout leaf development. Poor expression was observed at the vegetative stage of the plant (Figure 39).

In contrast, gene expression of cell wall invertase 6 (*Atcwinv6*, *6, 1-FEH*, AT5G11920) was found to be much lower than that of *Atcwinv3*. The highest levels of expression were observed in senescent leaves and roots from vegetative plants. This gene is moderately expressed, specifically in sepals, and at low levels in some development stages of the leaves. The lowest expression of the gene was identified in most stages of leaf development such as young and mature leaves, floral buds, shoot apical meristem and during seed development (Figure 39). Following the *in silico* gene expression patterns, quantitative PCR (qPCR) was carried out in order to confirm the expression pattern in those tissues where the highest FEH expression levels were predicted *in silico*.

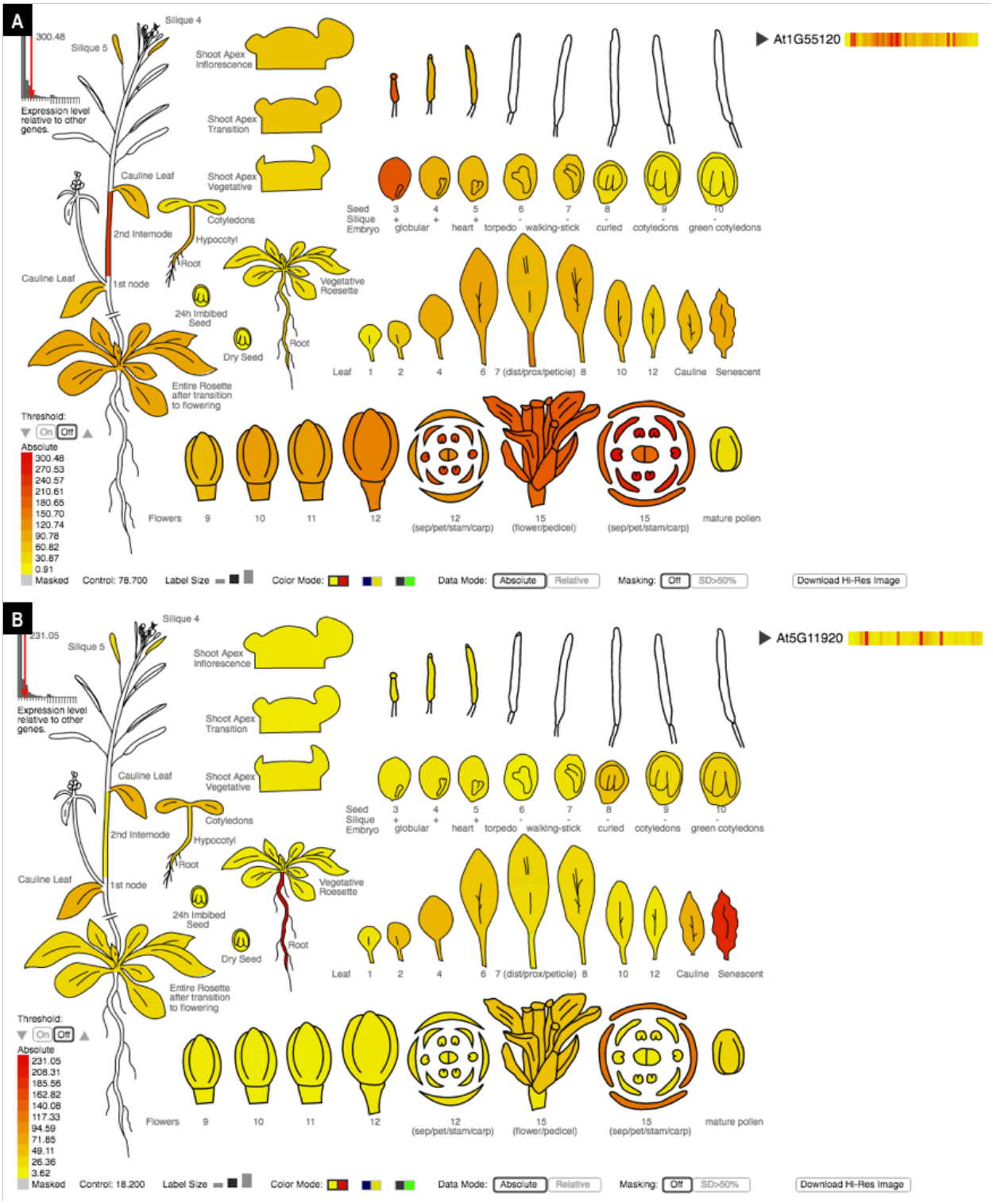


FIGURE 39. *In silico* gene expression of **A.** Cell wall invertase 3 (*Atcwinv3*, *6-FEH*, AT1G55120) and **B.** Cell wall invertase 6 (*Atcwinv6*, *6, 1-FEH*, AT5G11920) (bar.utoronto.ca, Schmid, *et al*., 2005 and Winter, *et al*., 2007).

qRT-PCR analysis of *Atcwinv3* (*6-FEH*) and *Atcwinv6* (*6, 1-FEH*) in WT Arabidopsis

Gene expression of *Atcwinv3* (*6-FEH*) based on qRT-PCR was consistent with that predicted in *silico*. A tendency to increased expression as leaves aged from plantlets, to young and later to mature leaves was observed, where the latter development stage showed highest gene expression. The lowest level of expression was observed in senescent leaves (tissue and development stage of reference). No differences were identified in roots from vegetative and reproductive plants, with expression only slightly higher in roots from vegetative plants. On the other hand, floral buds showed the second highest level of gene expression, whereas a decrease in expression was detected in fully opened flowers (Figure 40).

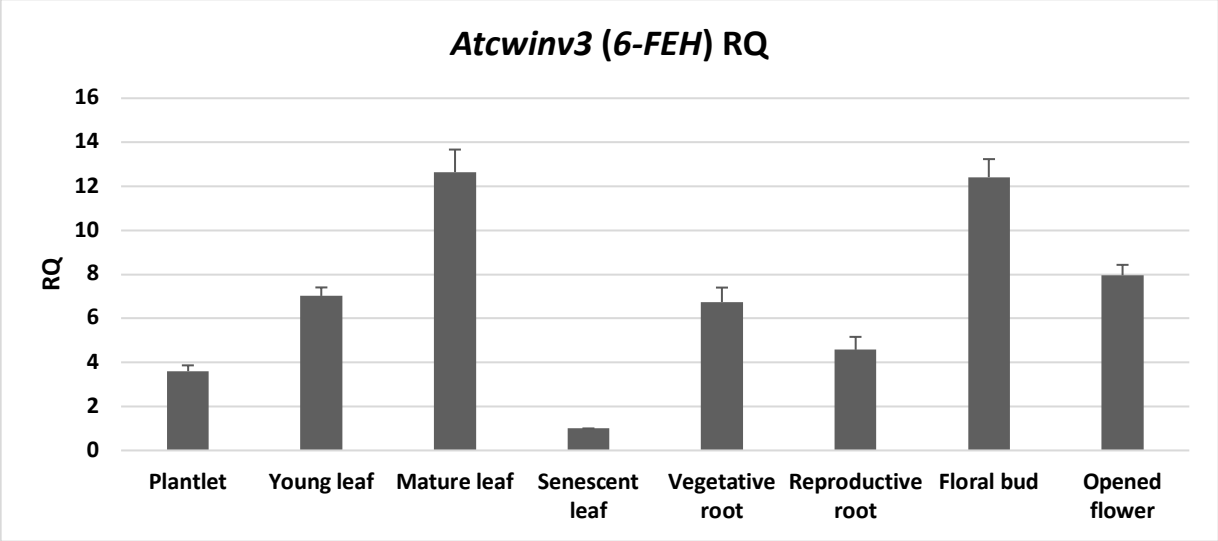


FIGURE 40. Relative gene expression quantification of *Atcwinv3* (*6-FEH*) from WT Arabidopsis tissues. RQ. Relative quantification. Senescent tissue was chosen as the quantitative reference data.

In contrast, expression of *Atcwinv6* (*6, 1-FEH*) was completely distinct from *Atcwinv3*. In this case, the reference tissue with lowest expression was the "plantlet". Gene expression in leaves showed a decreasing tendency throughout leaf development, achieving the

highest expression in young leaves and gradually decreasing until the senescent stage. This result was somehow unexpected given that the *in silico* gene expression prediction showed that senescent leaves were within the tissues with the highest gene expression levels (Figure 39). One notable difference between *Atcwinv 3* and *6*, was the gene expression levels detected in roots. *Atcwinv6* showed the highest expression in roots from vegetative plants while roots from reproductive plants showed almost a 50% reduction in gene expression. Floral tissues showed lowest gene expression both at the floral bud and the fully developed flower stages, showing no differences in gene expression throughout floral tissue development (Figure 41). The latter results were also consistent with those identified *in silico* (Figure 39).

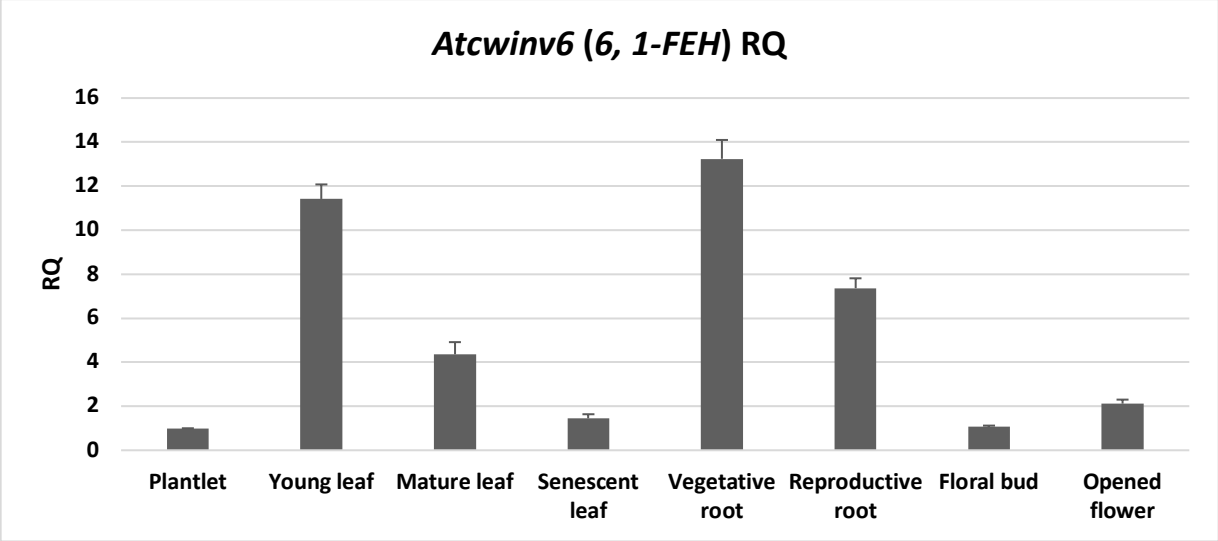


FIGURE 41. Relative gene expression quantification of *Atcwinv6* (*6, 1-FEH*) from WT Arabidopsis tissues. RQ. Relative quantification. Plantlet tissue was chosen as the quantitative reference data.

qRT-PCR analysis of *Atcwinv3* (*6-FEH*) and *Atcwinv6* (*6, 1-FEH*) in transgenic Arabidopsis lines

Endogenous FEH gene expression was also determined in the *p35S-Atq1SST-1* transformed lines. The results showed that the gene expression pattern of the two *FEH*

genes was also different in the *p35S-Atq1SST-1* Arabidopsis lines. Thus, the *Atcwinv3* (*6-FEH*) gene was shown to have the highest expression level in transgenic Arabidopsis line 5. This was also the transgenic line with highest *p35S-Atq1SST-1* expression. In general, *Atcwinv3* and *p35S-Atq1SST-1* showed similar patterns of fluctuating expression levels (Figure 33 and 42). On the other hand, *Atcwinv6* (*6, 1-FEH*) showed highest expression levels in lines 10, 15 and 17, followed by lines 22 and 30 while the rest of the lines showed the same low gene expression. No correlation between *Atcwinv6* and *p35S-Atq1SST-1* expression levels was observed (Figure 33 and 42).

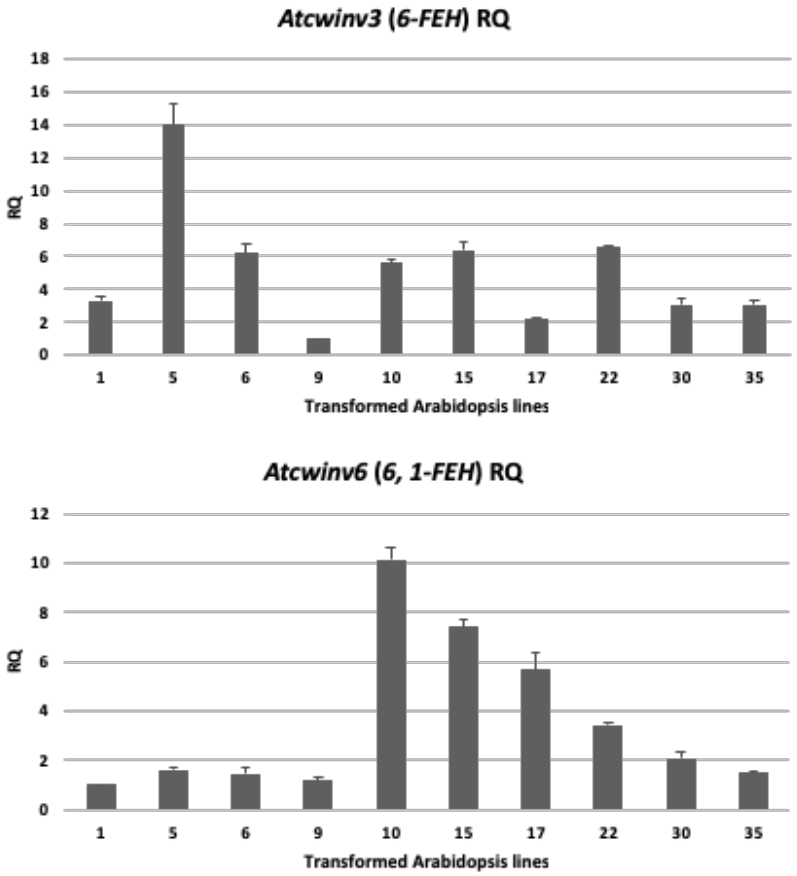


FIGURE 42. Relative gene expression quantification of *Atcwinv3* (*6-FEH*) and *Atcwinv6* (*6, 1-FEH*) in different transgenic Arabidopsis lines. RQ. Relative quantification, and the numbers represent the transgenic lines that were analyzed.

qRT-PCR analysis of *Atcwinv3* (*6-FEH*), *Atcwinv6* (*6, 1-FEH*) and the transgene (*Atq1SST-1)* in WT and transformed Arabidopsis lines

A direct comparison between *FEH* gene expression in WT and transgenic Arabidopsis was also carried out in order to confirm differences in expression levels in relation to the expression level of the *FEH* transcripts. For this experiment, only the most contrasting transgene lines were chosen (lines with the highest and lowest transgene expression levels, respectively).

By analyzing and comparing the expression levels of the genes in WT and transgenic lines, the results obtained previously were confirmed, since the expression of *Atcwinv3* was higher than *Atcwinv6* in mature leaves of WT Arabidopsis. Transgene expression levels were also consistent with higher expression levels in line 5 followed by lines 6, 35 and 1, respectively. Interestingly, the results showed that with the exception of Arabidopsis *p35S-Atq1SST-1* line 5, all transgenic lines showed a reduction of around 5 to 6-fold in the expression levels of both endogenous *FEH* genes (Figure 43). This suggests that ectopic expression of *p35S-Atq1SST-1* inhibits endogenous *FEH* gene expression. The *p35S-Atq1SST-1* line 5 showed very high levels of transgene expression and in this line, the specific expression of *Atcwinv3*, was increased in relation to the WT, indicating a differential response or function for each *FEH* gene.

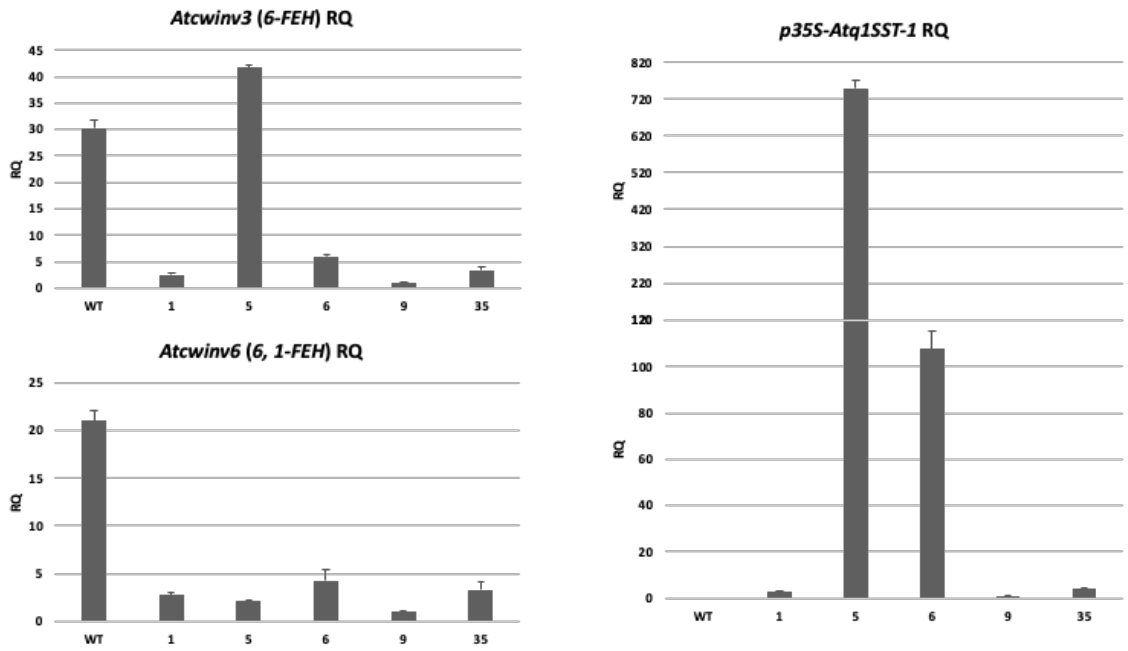


FIGURE 43. Relative gene expression quantification of the *p35S-Atq1SST-1* transgene, *Atcwinv3* (*6-FEH*) and *Atcwinv6* (*6, 1-FEH*) in transgenic lines compared to WT Arabidopsis. **RQ**. Relative quantification, WT. Wild type Arabidopsis and 1, 5, 6, 9 and 35. Code numbers of transformed Arabidopsis lines.

Discussion

Atq1SST-1 activity had previously been characterized in the heterologous system of *P. pastoris* (Avila-Fernández, *et al.,* 2007). However, limitations in methodology for transformation of *A. tequilana* have not yet permitted a functional characterization of the *Atq1SST-1* gene in agave. The model system of Arabidopsis presents an opportunity to functionally characterize *Atq1SST-1* in a plant species and to begin to determine the possible roles of the endogenous Arabidopsis *FEH* genes (*Atcwinv3* and *Atcwinv6*).

The generation of 35 transgenic Arabidopsis lines expressing the *Atq1SST-1* transgene was confirmed by both genomic and qRT-PCR analysis, and a selection of lines with varying levels of transgene expression were chosen for further analysis. Observation of phenotypes during the life-cycle of the transgenic lines in comparison to the WT reference uncovered little variation in size, form or number of leaves. However, on reaching the reproductive stage, transgenic lines produced shorter and fewer inflorescences in comparison to the WT reference.

Although transgene insertion and transcription were confirmed in transgenic lines, it was not completely possible to detect 1-kestose or FOS with a higher DP presence/synthesis in any of the transgenic lines by TLC or HPAEC-PAD. Nevertheless, it was confirmed by HPAEC-PAD that the carbohydrate profile was quite different between transgenic lines and the WT reference, showing drastic differences in fructose, glucose and sucrose as well as the presence/synthesis of possible maltose peaks in the transgenic samples, suggesting that carbohydrate metabolism has been altered in the transgenic lines. The latter could be, somehow, directly related to the *Atq1SST-1* transgene expression. However, without evidence of the presence of 1-kestose molecules, the accurate function of this enzyme in Arabidopsis remains questionable. Therefore, it will be necessary to further analyze generations of homozygous transgenic plants and improve on the water-soluble

carbohydrate extraction method, to determine, unambiguously, the presence or absence of fructan molecules.

Arabidopsis does not naturally synthesize fructans, but it has been reported that a vacuolar invertase could synthesize fructans *in vitro* under a high concentration of sucrose (De Coninck, *et al.,* 2005). Intriguingly, two genes encoding FEH type enzymes have also been characterized and are thought to participate in pathogen responses (De Coninck, *et al.,* 2005). Analysis of *in silico* and qRT-PCR expression patterns for the *FEH* genes uncovered distinct patterns of expression in different tissues and at different developmental stages. In order to determine the effect of ectopic expression of *Atq1SST-1* on the expression of these *FEH* genes, a qRT-PCR analysis of *Atq1SST-1*, and both *FEH* genes, was performed in the transgenic lines and compared to WT expression levels in leaves. Unexpectedly the expression of *Atq1SST-1* repressed rather than induced the expression of both *FEH* genes in all transgenic lines, with the exception of the *Atq1SST-1* line 5, which shows strong *Atq1SST-1* transgene expression levels and higher *Atcwinv3* gene expression than the WT reference. One explanation for these observations is that any 1-kestose in Arabidopsis plants is quickly processed by FEH activity of the endogenous Arabidopsis *FEH* genes and therefore not detectable. The increase in sucrose and fructose produced by this activity may lead to down regulation of *FEH* expression. In the case of *Atq1SST-1* line 5, higher levels of expression and putative synthesis of 1-kestose may stimulate *Atcwinv3* expression in order to process these molecules. The age/developmental stage of the plants that were tested for the presence of 1-kestose may also be a factor and transgenic lines should be analyzed at different points of the growth cycle to address this. The presence of high levels of maltose in the *Atq1SST-1* transgenic lines indicates that even though 1-kestose could not be detected, normal sugar (starch) metabolism is disrupted.

Additionally, morphological effects most obvious in floral tissue where in the wild type *Atcwinv3* is strongly expressed were observed as a result of transgene expression. This

may suggest that the morphological changes observed were due to repression of *FEH* (*Atcwinv3)* gene expression rather than directly due to the ectopic expression of *Atq1SST-1* and/or the putative presence of 1-kestose. Interestingly, only *Atcwinv3* responded to high levels of *Atq1SST-1* expression supporting the suggestion of functional differences between these two *FEH* genes based on their differential expression patterns. However, the basis for increased *Atcwinv3* expression in relation to strong expression of *Atq1SST-1* is unclear.

More extensive analysis in different tissues and developmental stages of the transgenic plants are necessary in order to begin examining the roles of the *FEH* genes in Arabidopsis in relation to *Atq1SST-1* expression. In addition, it will also be interesting to analyze the expression and activity of the other endogenous members of PGHF32 (vacuolar and cell wall invertases) in the transgenic Arabidopsis lines. The presence of maltose indicates putative effects on starch metabolism and it will also be interesting to determine whether changes occur in β-amylase expression or in the expression of other starch metabolism associated genes in the transgenic lines.

Conclusions

1. The cDNA encoding the FT *Atq1SST-1* isoform from *A. tequilana* was successfully cloned into an expression vector and expressed in the heterologous plant model Arabidopsis.

2. It was not possible to confirm the functional activity of the enzyme *in planta* because the expected 1-kestose metabolite synthesized by this enzyme was not detected in the transgenic extracts analyzed by chromatographic techniques.

3. Expression of the *Atq1SST-1* transgene repressed the expression of the endogenous Arabidopsis *FEH* genes leading to morphological changes in the floral structures of transgenic lines.

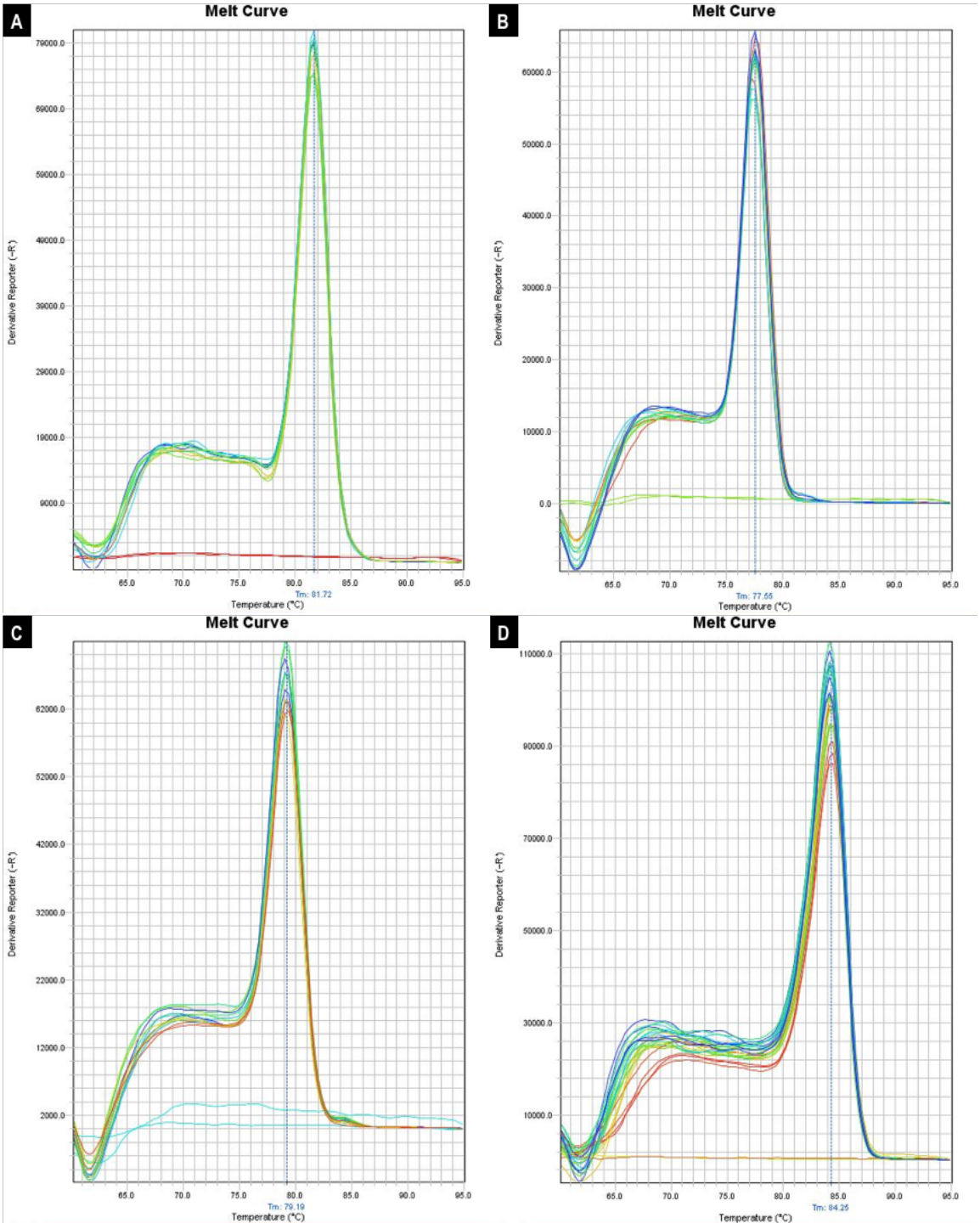
4. Strong transgene expression lead to the specific induction of *Atcwinv3*, suggesting functional differences between both Arabidopsis *FEH* genes.

5. The presence of maltose in transformed Arabidopsis might be related to a putative effect on carbohydrate metabolism (starch) by Atq1SST-1 transgene expression.

Perspectives

1. To obtain homozygous transgenic *Atq1SST-1* lines and complete the gene expression analysis of PGHF32 in different tissues and throughout the Arabidopsis life-cycle.
2. To analyze and to corroborate the functional activity of this protein *in planta* by carrying out protein purification.
3. To quantify in detail, and in different tissues, carbohydrates such as sucrose, starch and 1-kestose in transformed and WT Arabidopsis.
4. To carry out more detailed phenotypic analysis including roots, seeds, etc.
5. If fructans are present, perform abiotic stress analysis (e. g., drought and cold) in WT and transgenic Arabidopsis lines to find out if the presence of possible fructan molecules can increase plant tolerance against abiotic stress.
6. If fructans are present, perform biotic stress analysis by levan producing pathogenic bacteria such as certain pathovars of *Pseudomonas* or *Erwinia*, to confirm changes in the gene expression patterns of FEH genes, in order to support the role of these enzymes in plant defense.
7. To analyze the catalytic activity of FEH enzymes in WT and transformed Arabidopsis plants and compare it with the transcriptional results obtained from qPCR.
8. To investigate the effects on starch metabolism of *Atq1SST-1* expression in transformed Arabidopsis lines.

Appendix 1



Melt curves of PCR products of **A**. *Ubiquitin 11*, **B.** *Cell wall invertase 3* (*6-FEH*), **C.** *Cell wall invertase 6* (*6, 1-FEH*) and **D.** *AtqSST-1* in Arabidopsis.

References

1. Abraham-Juárez M. J., Martínez-Hernández A., Leyva-González M. A., Herrera-Estrella L. and Simpson J. (2010). Class I KNOX genes are associated with organogenesis during bulbil formation in Agave tequilana. Journal of Experimental Botany, 61, 4055-4067.

2. Arizaga S. and Ezcurra E. (1995). Insurance against reproductive failure in a semelparous plant: bulbil formation in *Agave macroacantha* flowering stalks. Oecologia, 101, 329-334.

3. Ávila de Dios, E., Gomez Vargas, A. D., Damián Santos, M. L. and Simpson, J. (2015). New insights into plant glycoside hydrolase family 32 in Agave species. Frontiers in Plant Science, 6, 594.

4. Ávila-Fernández A., Olvera-Carranza C., Rudiño-Piñera E., Cassab I. G., Nieto-Sotelo J. and López Munguía A. (2007). Molecular characterization of sucrose: sucrose 1-fructosyltransferase (1-SST) from *Agave tequilana* Weber var. azul. Plant Science, 173, 478-486.

5. Banguela A. and Hernández L. (2006). Fructans: from natural sources to transgenic plants. Biotecnología Aplicada, 23, 202-210.

6. Bielesky R. L. (1993). Fructan hydrolysis drives petal expansion in the ephemeral daylily flower. Plant Physiology, 103, 213-219.

7. Bonnett G. D., Sims I. M., Simpson R. J. and Cairns A. J. (1997). Structural diversity of fructan in relation to the taxonomy of the Poaceae. New Phytologist, 136, 11-17.

8. Borland, A. M., Hartwell, J., Weston, D. J., Schlauch, K. A., Tschaplinski, T. J., Tuskan, G. A. and Cushman, J. C. (2014). Engineering crassulacean acid metabolism to improve water-use efficiency. Trends in Plant Science, 19, 327-338.

9. Cimini S., Di Paola L., Giuliani A., Ridolfi A and De Gara L. (2016). GH32 family activity: a topological approach through protein contact networks. Plant Molecular Biology, 92, 401-10.

10. Cimini S., Locato V., Vergauwen R., Paradiso A., Cecchini C., Vendenpoel L., Verspreet J., Courtin C. M., D'Egidio M. G., Van den Ende W. and De Gara L. (2015). Fructan biosynthesis and degradation as part of plant metabolism controlling sugar fluxes during durum wheat kernel maturation. Frontiers in Plant Science, Plant Physiology, 6, 89.

11. Clough S. J. and Bent A. F. (1998). Floral dip: a simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana*. The Plant Journal, 16, 735-743.

12. Cortés-Romero, C., Martínez-Hernández, A., Mellado-Mojica, E., López, M. G. and Simpson, J. (2012). Molecular and functional characterization of novel fructosyltransferases and invertases from *Agave tequilana*. PloS One, 7, e35878.

13. De Coninck, B., Le Roy, K., Francis, I., Clerens, S., Vergauwen, R., Halliday, A. M., Smith S. M., Van Laere A. and Van den Ende, W. (2005). Arabidopsis AtcwINV3 and 6 are not invertases but are fructan exohydrolases (FEHs) with different substrate specificities. Plant, Cell and Environment, 28, 432-443.

14. Diario Oficial de la Federación, DOF. Retrieved September 18th, 2018 http://www.dof.gob.mx/nota_detalle.php?codigo=5282165&fecha=13/12/2012

15. Duchateau N., Bortlik K., Simmen U., Wiemken A. and Bancal P. (2005). Sucrose: fructan 6-fructosyltransferase, a key enzyme for diverting carbon from sucrose to fructans in barley leaves. Plant Physiology, 107, 1249-1255.

16. Edelman J. and Jefford T. G. (1968). The mechanisim of fructosan metabolism in higher plants as exemplified in *Helianthus tuberosus*. New Phytologist, 67, 517-531.

17. Escobar-Guzmán R. E., Zamudio Hernández F., Gil Vega K., and Simpson J. (2008). Seed production and gametophyte formation in *Agave tequilana* and *Agave americana*. Botany, 86, 1343-1353.

18. Fujishima M., Sakai H., Ueno K., Takahashi, N., Onodera S., Benkeblia N. and Shiomi N. (2005). Purification and characterization of a fructosyltransferase from onion bulbs and its key role in the synthesis of fructooligosaccharides *in vivo*. New Phytologist, 165, 513-524.

19. García-Mendoza A. and Galván V. R. (1995). Riqueza de las familias Agavaceae y Nolinaceae en México. Botanical Sciences, 56, 7-24.

20. García-Mendoza A. J. (2007). Los Agaves de México. Red de revistas científicas de América Latina y El Caribe, España y Portugal, 087, 14-23.

21. García-Mendoza, A. J. (2002). Distribution of the genus *Agave* (Agavaceae) and its endemic species in Mexico. Cactus and Succulent Journal, 74, 177-187.

22. Gentry J. S. (1982). Agaves of Continental North America. University of Arizona Press, Tucson, AZ, USA.

23. Hendry, G. A. F. (1993). Evolutionary origins and natural functions of fructans. A climatological, biogeographic and mechanistic appraisal. New Phytologist, 123, 3-14.

24. Holtum, J. A. M., Chambers D., Morgan T. and Tan D. K. Y. (2011). *Agave* as a biofuel feddstok in Australia. GCB Bioenergy, 3, 58-67.

25. Irish M. and Irish G. (2000). Agaves, yuccas and related plants. A gardener guide. Timbers Press. Portland, Oregon, USA.

26. Karimi M., Inze D. and Depicker A. (2002). Gateway vectors for *Agrobacterium*-mediated plant transformation. Trends in Plant Science, 7, 193-195.

27. Kawakami A., Yoshida M. and Van den Ende W. (2005). Molecular cloning and functional analysis of a novel 6&1-FEH from wheat (*Triticum aestivum* L.) preferentially degrading small branched graminans like bifurcose. Gene, 358, 93-101.

28. Koch K. (1996). Carbohydrate-modulated gene expression in plants. Annual Review of Plant Physiology and Plant Molecular Biology, 47, 509-540.

29. Lammens W., Roy K. L., Yuan S., Vergauwen R., Rabijns A., Van Laere A., Strelkov S. V. and Van Den Ende W. (2012). Crystal structure of 6-SST/6-SFT from *Pachysandra terminalis*, a plant fructan biosynthesizing enzyme in complex with its acceptor substrate 6-kestose. The Plant Journal, 70, 205-219.

30. Livingston D. P. and Henson C. A. (1998). Apoplastic sugars, fructans, fructan exohydrolase, and invertase in winter oat: responses to second phase cold hardening. Plant Physiology, 116, 403-408.

31. Livingston D. P. I., Chatterton N. J. and Harrison P. A. (1993). Structure and quantity of fructan oligomers in oat (*Avena* spp.). New Phytologist, 123, 725-34.

32. Livingston D. P., Hincha D. K. and Heyer A. G. (2009). Fructan and its relationship to abiotic stress tolerance in plants. Cellular and Molecular Life Sciences, 66, 2007-2023.

33. Livingston III D. P., Hincha D. K. and Heyer A. G. (2007). The relationship of fructan to abiotic stress tolerance in plants. Recent Advances in Fructooligosaccharides Research, 181-199.

34. López M. G., Mancilla-Margalli N. A. and Mendoza-Díaz G. (2003). Molecular structures of fructans from *Agave tequilana* Weber var. azul. Journal of Agricultural and Food Chemistry, 51, 7835-7840.

35. Lüscher M. and Nelson C. J. (1995). Fructosyl transferase activities in the leaf growth zone of tall fescue. Plant Physiology, 107, 1419-1425.

36. Mancilla-Margalli N. A. and López M. G. (2006). Water-soluble carbohydrates and fructan structure patterns from *Agave* and *Dasylirion* species. Journal of Agricultural and Food Chemistry, 54, 7832-7839.

37. Mellado-Mojica E. and López, M.G. (2012). Fructan metabolism in *A. tequilana* Weber Blue variety along its developmental cycle in the field. Journal of Agricultural and Food Chemistry, 60, 11704-11713.

38. Mellado-Mojica E., López-Medina T. L. and López M. G. (2009). Developmental variation in *Agave tequilana* Weber var. azul stem carbohydrates. Dynamic Biochemistry, Process Biotechnology and Molecular Biology, 3, 34-39.

39. Naumoff D. G. (2001). β-Fructosidase superfamily: homology with some α-Larabinases and β-D-xylanases. Proteins, 42, 66-76.

40. Nava-Cruz N. Y., Medina-Morales A., Martínez J. L., Rodríguez R. and Aguilar C. N. (2015). *Agave* biotechnology: an overview. Critical Reviews in Biotechnology, 35, 546-559.

41. Nilsson U. and Dahlqvist A. (1986). Cereal fructosans: part 2. Characterization and structure of wheat fructosans. Food Chemistry, 22, 95-106.

42. Pavis N., Boucaud J. and Prud'Homme M. P. (2001). Fructans and fructan-metabolizing enzymes in leaves of *Lolium perenne*. New Phytologist, 150, 97-109.

43. Pollock C. J. (1986). Environmental effects on sucrose and fructan metabolism. Plant biochemistry and physiology, 5, pp 32-46.

44. Pollock C. J., Eagles C. F. and Sims I. M. (1988). Effect of photoperiod and irradiance changes upon development of freezing tolerance and accumulation of soluble carbohydrate in seedlings of *Lolium perenne* grown at 2 °C. Annals of Botany, 62, 95-100.

45. Ritsema T., Joling J. and Smeekens J. (2003). Patterns of fructan synthesized by onion fructan: fructan 6G-fructosyltransferase expressed in tobacco BY2 cells: is fructan: fructan 1-fructosyl transferase needed in onion? New Phytologist; 160, 61-67.

46. Schmid M., Davison T. S., Henz S. R., Pape U. J., Demar M., Vingron M., Schölkopf B., Weigel D. and Lohmann J. U. (2005). A gene expression map of *Arabidopsis thaliana* development. Nature Genetics, 37, 501-506.

47. Secretaría de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación. (2017). Agave tequilero y mezcalero mexicano. México.

48. Sprenger N., Schellenbaum L., Van Dun K., Boller T. and Wiemken A. (1997). Fructan synthesis in transgenic tobacco and chicory plants expressing barley sucrose: fructan 6-fructosyl transferase. FEBS Letters, 400, 355-358.

49. Sprenger, N., Bortlik, K., Brandt, A., Boller, T. and Wiemken, A. (1995). Purification, cloning and functional expression of sucrose: fructan 6-fructosyltransferase, a key enzyme of fructan synthesis in barley. Proceeding of the National Academy of Sciences of the United Stated of America, 92, 11652-11656.

50. Stoop, J. M., Van Arkel, J., Hakkert, J. C., Tyree, C., Caimi, P. G., and Koops, A. J. (2007). Developmental modulation of inulin accumulation in storage organs of transgenic maize and transgenic potato. Plant Science, 173, 172-181.

51. Suzuki M. and Nass H. G. (1988). Fructan in winter wheat, triticale, and fall rye cultivars of varying cold hardiness. Canadian Journal of Botany, 66, 1723-1728.

52. Tarkowski Ł. P. and Van den Ende W. (2015). Cold tolerance triggered by soluble sugars: a multifaceted countermeasure. Frontiers in Plant Science, Plant physiology, 6, 203.

53. Tauzin A. S. and Giardina T. (2014). Sucrose and invertases, a part of the plant defense response to the biotic stresses. Frontiers in Plant Science, Plant-Microbe Interaction, 5, 293.

54. Tymowska-Lananne Z. and Kreis M. (1998). The plant invertases: physiology, biochemistry and molecular biology. Advances in Botanical Research, 28, 71-117.

55. Valluru R. and Van den Ende W. (2008). Plant fructans in stress environments: emerging concepts and future prospects. Journal of Experimental Botany, 59, 2905-2916.

56. van Arkel J., Séveniera R., Hakkerta J. C., Bouwmeesterb H. J., Koopsa A. J. and van der Meer I. M. (2013). Tailor-made fructan synthesis in plants: A review. Carbohydrate Polymers, 93, 48-56.

57. Van den Ende W. (2013). Multifunctional fructans and raffinose family oligosaccharides. Frontiers in Plant Science, Plant Physiology, 4, 247.

58. Van den Ende W., Clerens S., Vergauwen R., Boogaerts D., Le Roy K., Arckens L. and Van Laere A. (2006). Cloning and functional analysis of a high DP fructan: fructan 1-fructosyl transferase from *Echinops ritro* (Asteraceae): comparison of the native and recombinant enzymes. Journal of Experimental Botany, 57, 775-789.

59. Van den Ende W., De Coninck B., Clerens S., Vergauwen R. and Van Laere A. (2003). Unexpected presence of fructan 6-exohydrolases (6-FEHs) in non-fructan plants: characterization, cloning, mass mapping and functional analysis of a novel 'cell-wall invertase-like' specific 6-FEH from sugar beet (*Beta vulgaris* L.). The Plant Journal, 36, 697-710.

60. Van den Ende W., Yoshida M., Clerens S., Vergauwen R. and Kawakami A. (2005). Cloning, characterization and fuctional analysis of novel 6-kestose exohydrolases (6-KEHs) from wheat (*Triticum aestivum*). New Phytologist, 166, 917-932.

61. Van der Meer I. M., Koops A. J., Hakkert J. C. and Van Tunen A. J. (1998). Cloning of the fructan biosynthesis pathway of Jerusalem artichoke. The Plant Journal, 15, 489-500.

62. Vergauwen R., Van den Ende W. and Van Leare A. (2000). The role of fructan in flowering of *Campanula rapunculoides*. Journal of Experimental Botany, 51, 1261-1266.

63. Verhaest M., Ende W. V., Roy K. L., De Ranter C. J., Van Laere A. and Rabijns A. (2005). X-ray diffraction structure of a plant glycosyl hydrolase family 32 protein: fructan 1-exohydrolase IIa of *Cichorium intybus*. The Plant Journal, 41, 400-411.

64. Verhaest M., Lammens W., Le Roy K., De Coninck B., De Ranter C. J., Van Laere A., Van Den Ende W. and Rabijns A. (2006). X-ray diffraction structure of a cell-wall invertase from *Arabidopsis thaliana*. Acta Crystallographica, 62, 1555-1563.

65. Verhaest M., Lammens W., Le Roy K., De Ranter C. J., Van Laere A., Rabijns A. and Van den Ende W. (2007). Insights into the fine architecture of the active site of chicory fructan 1-exohydrolase: 1-kestose as substrate vs sucrose as inhibitor. New Phytologist, 174, 90-100.

66. Versluys M., Kirtel O., Toksoy Öner E. and Van den Ende W. (2018). The fructan syndrome: Evolutionary aspects and common themes among plants and microbes. Plant Cell Environ, 41, 16-38.

67. Versluys M., Tarkowski Ł. P. and Van den Ende W. (2017) Fructans as DAMPs or MAMPs: evolutionary prospects, cross-tolerance, and multistress resistance potential. Frontiers in Plant Science, 7, 2061.

68. Verspreet J., Cimini S., Vergauwen R., Dornez E., Locato V., LeRoy K., De Gara L, Van den Ende W, Delcour J. A. and Courtin C. M. (2013). Fructan metabolism in developing wheat (*Triticum aestivum* L.) kernels. Plant and Cell Physiology, 54, 2047-2057.

69. Vijn I. and Smeekens S. (1999). Fructan: More than a reserve carbohydrate?. Plant Physiology, 120, 351-359.

70. Vijn I., Van Dijken A., Sprenger N., Van Dun K., Weisbeek P., Wiemken A. and Smeekens S. (1997). Fructan of the inulin neoseries is synthesized in transgenic chicory plants (*Chichorium intybus* L.) harbouring onion (*Allium cepa* L.) fructan: fructan 6G-fructosyltransferase. The Plant Journal, 11, 387-398.

71. Wei J. Z. and Chatterton N. J. (2001). Fructan biosynthesis and fructosyltransferase evolution: Expression of the 6-SFT (sucrose: fructan 6- fructosyltransferase) gene in crested wheatgrass (*Agropyron cristatum*). Journal of Plant Physiology, 158, 1203-1213.

72. Wei J. Z., Chatterton N. J., Harrison P. A., Wang R. R. and Larson S. R. (2002). Characterization of fructan biosynthesis in big bluegrass (*Poa secunda*). Journal of Plant Physiology, 159, 705-715.

73. Weschke W., Panitz R., Gubatz S., Wangy Q., Radchuk R., Weber H. and Wobus U. (2003). The role of invertases and hexose transporters in controlling sugar ratios in maternal and filial tissues of barley caryopses during early development. The Plant Journal, 33, 395-411.

74. Weyens G., Ritsema T., Van Dun K., Meyer D., Lommel M., Lathouwers J, Rosquin I, Denys P., Tossens A., Nijs M., Turk S., Gerrits N., Bink S., Walraven B., Lefebvre M. and Smeekens S. (2004). Production of tailor-made fructans in sugar beet by expression of onion fructosyltransferase genes. Plant Biotechnology Journal, 2, 321-327.

75. Winter D., Vinegar B., Nahal H., Ammar R., Wilson G. V. and Provart N. J. (2007). An “electronic fluorescent pictograph” browser for exploring and analyzing large-scale biological data sets. PLoS ONE, 2, 718.

76. Yildiz S. (2011). The metabolism of fructooligosaccharides and fructooligosaccharide-related compounds in plants. Food Reviews International, 27, 16-50.