Spring 5-9-2010

Characterization of the Putative Xyloglucan Glycosyltransferase GT14 in Arabidopsis thaliana

Najam R. Syed

University of Connecticut - Storrs, nsyed88@gmail.com

Honors Scholar Thesis

Characterization of the Putative Xyloglucan Glycosyltransferase GT14
in *Arabidopsis thaliana*

Najam Syed

University of Connecticut
Department of Molecular and Cell Biology
Spring 2010

Thesis Advisor: Wolf-Dieter Reiter, PhD
Honors Advisor: Johann Peter Gogarten, PhD

Abstract

Plant cell walls largely consist of matrix polysaccharides that are linked to cellulose microfibrils. Xyloglucan, the primary hemicellulose of the cell wall matrix, consists of a repeating glucose tetramer structure with xylose residues attached to the first three units ('XXXG'). In *Arabidopsis thaliana*, the core XXXG structure is further modified by enzymatic addition of galactose and fucose residues to the xylose side chains to produce XLXG, XXLG, XLLG and XLFG structures. GT14 is a putative glycosyltransferase in the GT47 gene family. Initial predictions of GT14's hydrophobic regions, based on its translated amino acid sequence, are almost identical to its *Arabidopsis* homolog MUR3, which is a xyloglucan galactosyltransferase targeted to the Golgi membrane. This suggests that, like MUR3, GT14 possesses a transmembrane domain and that it is targeted to the Golgi.

The monosaccharide composition of leaves from T-DNA insertion knockouts of *GT14* was analyzed by gas-liquid chromatography. The *gt14* plants were found to have lower fucose and higher mannose contents than wild type plants. Analysis of cell wall and soluble fractions from *gt14* and wild type plants revealed that most of the deficiency in fucose was accounted for in the cell wall, supporting the idea that GT14's target is xyloglucan. Finally, *gt14* and wild type plants were transformed with *GT14* for complementation and overexpression analysis. The majority of transformed plants did not show significant changes with regard to monosaccharide composition. This may be because the plants were in the T1 generation and, thus, hemizygous. Analysis of homozygous plants in the T2 generation may reveal noticeable changes.

Further studies on the xyloglucan composition of *gt14* plants are necessary to put the observed reduction in cell wall fucose into a meaningful context.

Acknowledgments

Table of Contents

List of Figures and Tables

Introduction

i. The plant cell wall

  The plant cell wall is composed of cellulose microfibrils that are interconnected with hemicelluloses in an extracellular matrix of pectic polysaccharides. The cell wall is involved in regulating growth of the plant, as well as providing structural integrity for the cell and plant body (Carpita and Gibeaut, 1993). This serves to provide a framework for a number of other functions vital to the plant's development.

  Plants have two types of cell walls: primary and secondary. The primary cell wall is deposited while the plant is growing; thus, it must confer mechanical strength without sacrificing pliability in order to allow cell expansion (Reiter, 2002). The primary cell wall is principally composed of networks of three distinct polysaccharides: cellulose, hemicelluloses and pectins (McNeil et al., 1984; Bacic et al., 1988; McCann and Roberts, 1991). Cellulose is a 1,4-β-D-glucan synthesized at the plasma membrane (Reiter, 2002). It is organized into a microfibril structure containing 36 parallel, paracrystalline polysaccharide chains (Delmer, 1999). These are bound together by the hemicelluloses and pectins (i.e., matrix polysaccharides), which are synthesized in the Golgi cisternae (Reiter, 2002). Secondary cell walls are, as the name implies, synthesized after the formation of the primary (growing) cell wall. In *Arabidopsis thaliana* and other plants, some specialized cell types, such as the xylem and interfascicular fibers, possess a secondary cell wall (Turner and Somerville, 1997).

ii. Xyloglucan in the plant cell wall
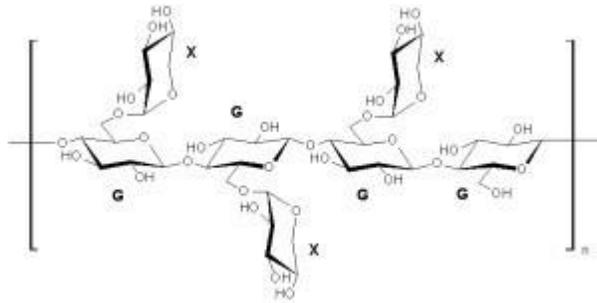
Figure 1. A representative unit of the hemicellulose xyloglucan. G = glucose, X = xylose.

The most abundant hemicellulose found in the primary cell wall of most higher plants is xyloglucan. Xyloglucan's structure usually consists of a 1,4-β-D-glucan backbone with 1,6-α-D-xylose moieties linked to the first three glucose residues. This core xyloglucan oligomer is notated 'XXXG'. A number of glycosyltransferases are responsible for modification of this initial building block. A D-galactose residue may be added to the second and third xylose residues with a β-1,2-linkage, resulting in the possible xyloglucan structures XLXG, XXLG, and XLLG. An L-fucose residue may be added at the second galactose residue in an α-1,2-linkage to form XLFG or XXFG. Several hypotheses for the purpose of the α-L-Fuc-1,2-β-D-Gal-1,2 side chain have been proposed. Computer modeling has suggested that the disaccharide group straightens the D-glucan backbone in order to facilitate the formation of hydrogen bonds with cellulose microfibrils (Levy et al., 1991, 1997). Another hypothesis proposes that xyloglucan oligomers with the fucosylated side chain are involved in auxin-induced growth and expansion (York et al., 1984; Fry, 1994).

iii. Glycosyltransferases involved in the modification of xyloglucan

Characterizing glycosyltransferases involved in cell wall polysaccharide synthesis and modification has been difficult due not only to the large number of genes involved, but also their intricate interrelationships. One estimate suggests that up to 15% of genes in the plant genome may be involved with cell wall synthesis, remodeling and turnover (Carpita et al., 2001). The *Arabidopsis thaliana* genome itself has over 800 identifiable carbohydrate active enzymes (Brown et al., 2005). In other words, development and maintenance of the plant cell wall represents a significant proportion of the plant's resources. It is no surprise, then, that the functions of and interaction between genes that encode carbohydrate active enzymes are complex.

Chemically mutagenized *Arabidopsis* plants that were screened for an abnormal cell wall monosaccharide composition were found to have defects at 11 particular loci. These 11 loci, designated *mur1* through *mur11*, fell into three groups: 1) mutant lines with reductions in the amount of a single monosaccharide, 2) those with absence of a monosaccharide and 3) plants with changes in the relative amounts of multiple monosaccharides. In *mur1* plants, fucose was essentially absent (Reiter et al., 1993). *mur2* and *mur3* suffered a reduction in cell wall fucose content by roughly 50%, *mur4* through *mur7* showed reductions in arabinose, and *mur8* showed a reduction in rhamnose content. *mur9* through *mur11*, which fell into the third category, displayed complex changes in their monosaccharide composition (Reiter et al., 1997).

Among the mutations that reduced the amount of fucose, the *mur1* plants were deficient because of missense mutations in the *GMD2* gene. The outcome of the defect is the production of nonfunctional isoforms of the enzyme GDP-D-Man 4,6-dehydratase

(Freshour et al., 2003). GDP-D-Man 4,6-dehydratase is necessary for the de novo

synthesis of the activated form of L-fucose, GDP-L-fucose (Mulichak et al., 2002). The

mur1 plants displayed a dwarfed phenotype (Reiter et al., 1997), though the phenotype

can be rescued by growing the plants in the presence of high concentrations of borate

(O'Neill et al., 2001). Boron is responsible for cross-linking the pectic polysaccharide

rhamnogalacturonan II (Kobayashi, et al., 1996), and structural changes in the fucose-

deficient mutants reduce the ability of rhamnogalacturonan II to form cross-links.

Therefore, the mutants require a greater boron concentration than the wild type plants to

achieve normal cross-linking. It appears that borate-complexed rhamnogalacturonan II,

not fucosylated xyloglucan, is primarily responsible for mechanical strength in

Arabidopsis (Ryden et al., 2003). Regardless, the fucosylation of xyloglucan in *mur1*

plants was largely restored by growing the plants in fucose-supplemented media (Pauly et

al., 2001), lending credence to the idea that the *mur1* defect is in the de novo synthesis of

fucose. The *mur2* line was found to have a loss-of-function mutation in the

fucosyltransferase AtFUT1 (Vanzin et al., 2002). Because AtFUT1 is a

fucosyltransferase specific to xyloglucan, the *mur2* mutation only affects fucose in

xyloglucan and not the entire plant as with *mur1*. In *mur3* plants, galactosylation of the

third xylose residue of xyloglucan is absent (Madson et al., 2003). This, of course,

precludes the addition of any fucose to xyloglucan. Therefore, the total fucose content of

mur3 plants is roughly half that of their wild type counterparts. The proportion of

xyloglucan with a galactosylated central xylose residue (XLXG) is significantly higher in

the *mur3* plants, perhaps to compensate for the loss of the α-L-Fuc-1,2-β-D-Gal side chain

(Madson et al., 2003).

iv. *MUR3* homologs and *GT14*

The *Arabidopsis* genome was searched using the PSI-BLAST algorithm (Altschul et al., 1997) with the MUR3 protein as the query sequence. This identified 38 coding regions in the genome with a high degree of sequence similarity to *MUR3*. These proteins are part of glycosyltransferase family 47 (GT47) as defined by Henrissat and Davies (Henrissat and Davies, 2000; see http://www.cazy.org/). They share the pfam03016 structural motif, which also represents the glucuronosyltransferase domain of animal exostosins. Results of the PSI-BLAST search were used to divide the GT47 family into a number of subgroups. MUR3 and 10 related homologs (GT11 through GT20) were part of subgroup A (Li et al., 2004).

The focus of my honors thesis is the characterization of the MUR3 homolog GT14. Leaves from a mutant line of plants with a T-DNA insertion in the *GT14* gene were analyzed for changes in carbohydrate composition. Wild type and knockout plants were also transformed with *GT14* for complementation and overexpression analysis. Finally, GT14's amino acid sequence was used to predict its location and orientation in cellular membranes.

Materials and Methods

Insertion of *GT14* into pCAMBIA and pSPDK vectors

A 1.566 kb *GT14* cDNA, which was obtained from a T-DNA insertion mutant from the Salk Institute Genome Analysis Laboratory (La Jolla, CA), was amplified via PCR using the following primers:

Forward primer containing a KpnI restriction site:

5'-GCGCGGTACCATGCGACCCAAGAATTATTCTCAGATGGAG-3'

Reverse primer containing a BamHI restriction site:

5'-CCGCGGATCCTTAAGCCAAAACTTGAGGTTCTTCATATGG-3'

The following program was used for PCR amplification:

1) 1 min at 94°C (1 cycle)

2) 30 s at 94°C; ramp down to 55°C for 30 s; ramp up to 68°C for 2 min (25 cycles)

3) 10 min at 68°C; ramp down to and maintain at 4°C

The amplification product was isolated through gel electrophoresis and gel extraction. The gel was cast using TAE buffer, 0.8% agarose and 0.4 μg/ml ethidium bromide, and run at a field strength of approximately 10 V/cm. The PCR product was run alongside a 1 kb DNA ladder. The desired band was excised, and the slice was purified using a QIAquick Gel Extraction Kit following the instructions provided by the manufacturer. The gel slice was placed in a 2.0 ml microfuge tube and incubated at approximately 65°C until melted. 150 μl of isopropanol was added to the tube and the sample was transferred to a QIAquick spin column and centrifuged. Then, 0.5 ml Buffer QG was added and the tube was centrifuged. Next, 750 μl of wash buffer was added and

the sample was again centrifuged. The sample was eluted into a clean tube with 50 μl of water.

GT14 was inserted into the pCAMBIA1301 vector (Fig. 2), which contains a kanamycin resistance gene for bacterial selection and a hygromycin resistance gene for plant selection. The GT14 insert and the pCAMBIA vector were treated with the restriction enzymes BamHI and KpnI in a double digestion. Ligation was achieved using T4 DNA ligase by adding the insert and vector in a 3:1 ratio and leaving the mixture overnight at 16°C. The 13.4 kb ligated plasmid was isolated via gel extraction as described above.
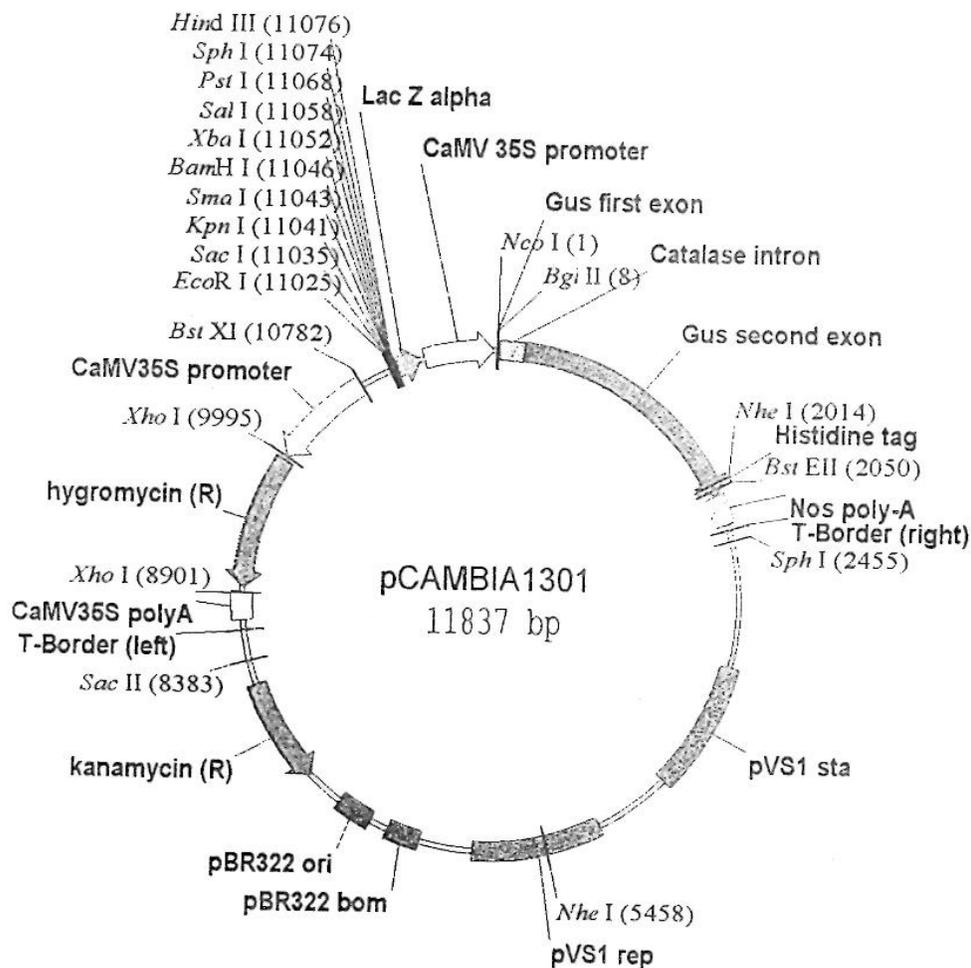
Figure 2. The pCAMBIA1301 vector and its multiple cloning sites.

GT14 was inserted into the pSPDK1677 vector (kindly provided by Dr. Dinesh-Kumar at Yale University; see Fig. 3B for the vector map and Appendix for the vector sequence) by means of the pENTR 1A dual selection entry vector (Fig. 3A and Fig. 4B), based on Invitrogen's Gateway cloning method. The method involves inserting the gene of interest (*GT14*) between the *attL* sites of the entry vector by performing a digestion and subsequent ligation. The *attL* sites in the entry vector, which flank the gene of interest, are recombined with *attR* sites in the destination vector's gateway cassette. The pSPDK vector contains a spectinomycin resistance gene for bacterial selection and a gentamicin resistance gene for plant selection.
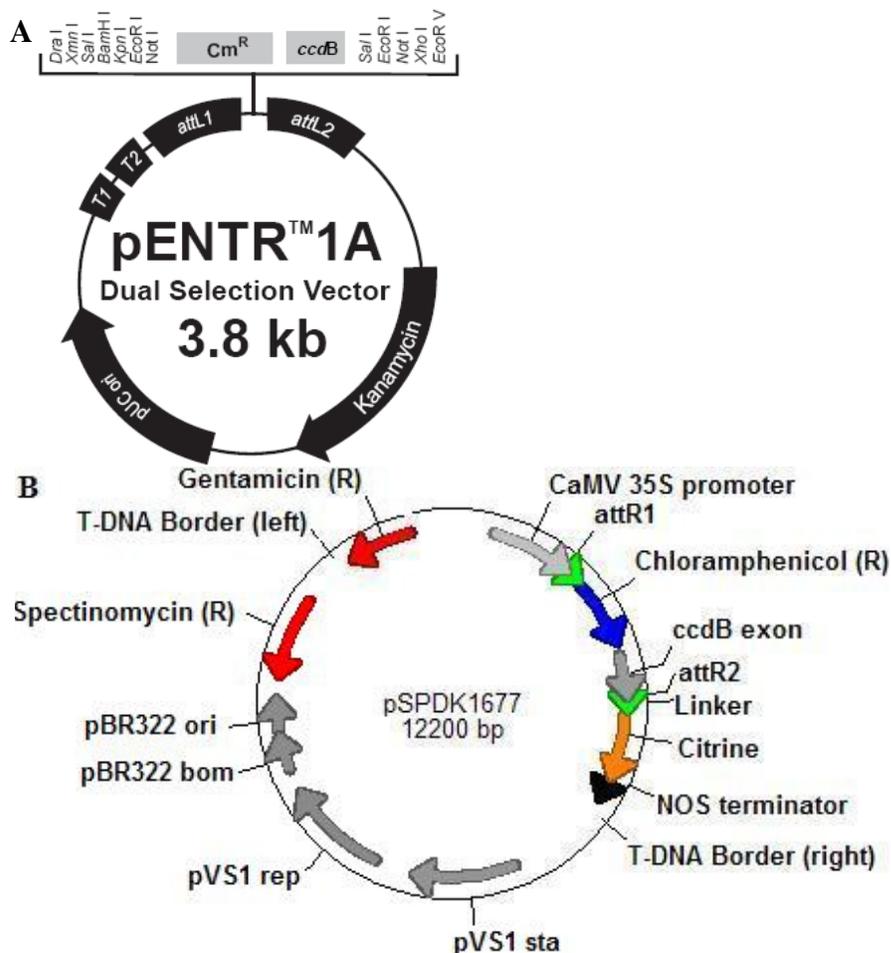
Figure 3. Maps of the pENTR 1A entry vector **(A)** and the pSPDK1677 destination vector **(B)**.

Primers for inserting *GT14* into pENTR 1A were designed as follows:

Forward primer containing a KpnI restriction site:

5'-GCGCGGTACCATGCGACCCAAGAATTATTCTCAGATGGAG-3'

Reverse primer containing a NotI restriction site:

5'-CCGCGCGGCCGCGAAGCCAAAACTTGAGGTTCTTCATATGGAGAA-3'

E. coli transformation and selection

100 μl of DH5α chemically competent cells were combined with 4 μl of ligated

vector (containing insert *GT14*) and 8 μl double distilled H_2O in 1.5 ml microfuge tubes

and left on ice for 30 minutes. The cells were exposed to a temperature of 42°C for 2

minutes and immediately returned to ice. The cells were allowed to grow at 37°C for 1-2

hours in SOC media (2% tryptone, 0.5% yeast extract, 10 mM sodium chloride, 2.5 mM

potassium chloride, 10 mM magnesium chloride, 10 mM magnesium sulfate, 20 mM

glucose). The mixture was centrifuged and all but 100-150 μl of the SOC media was

discarded. The pellet was resuspended, and the resuspended solution was distributed onto

a 2% LB agar plate (1% tryptone, 0.5% yeast extract, 1% sodium chloride, 2% agar) and

left overnight at 37°C. The LB agar plates contained either 50 μg/ml kanamycin, for cells

transformed with pCAMBIA, or 50 μg/ml spectinomycin, for cells transformed with

pSPDK.

Uptake of the entire plasmid was verified by performing a PCR colony screen of

bacteria that grew on the plate. Liquid LB media was inoculated with positive colonies

and incubated overnight at 37°C. Bacterial DNA from these cells was isolated using the

QIAGEN Plasmid Maxi Prep kit as per the instructions provided by the manufacturer.

The plasmid DNA was purified using gel electrophoresis. A small amount of this plasmid

DNA was double digested with the aforementioned restriction enzymes and run on the

gel to confirm the presence of the vector and insert bands (Fig. 4A).
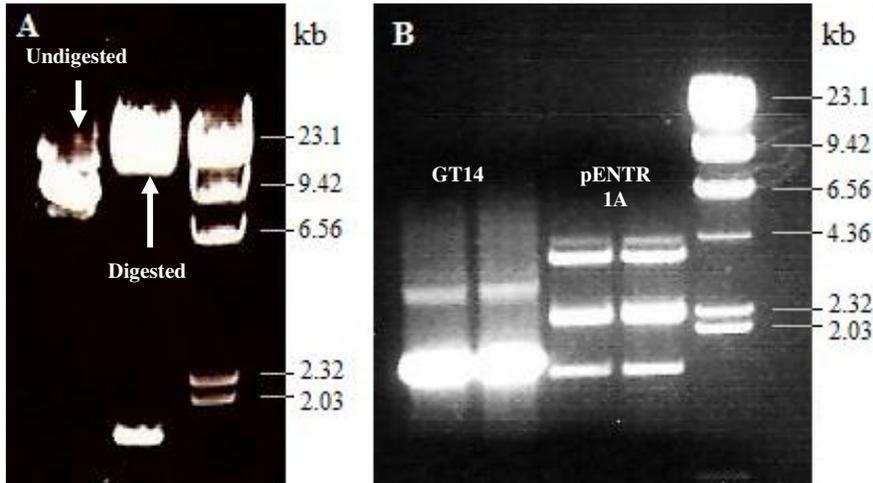
Figure 4. (A) Gel electrophoresis of the PCR amplification product of pCAMBIA (11.8 kb) with insert
GT14 (1.6 kb). One sample was treated with restriction enzymes KpnI and BamHI and run alongside an
undigested sample. **(B)** Gel electrophoresis of the PCR amplification products of *GT14* (1.6 kb) and
pENTR 1A vector (3.8 kb). The samples were treated with restriction enzymes KpnI and NotI. In the
pENTR 1A lanes, the 1.5 kb and 2.3 kb fragments represent the segments of the digested plasmid. The 3.8
kb fragment represents undigested material.
        Both sets of samples were run against HindIII-digested λ phage DNA as a size marker on a 0.8%
agarose gel.

The DNA was sequenced in a Ready-to-Run reaction. The following internal sequence

primers were designed for this purpose:

Forward primer: 5'-CGGTGTTTTGTCTTCAGCCACCGG-3'

Reverse primer: 5'-GGTCTCTACCAGACATTCGACCCC-3'

        Strains of *E. coli* that had been successfully transformed with pCAMBIA and

GT14-pCAMBIA were stored at -80°C in 15% glycerol.

Agrobacterium tumefaciens GV3101 transformation

Approximately 5 μl of DNA was added to 50-100 μl of chemically competent cells of *A. tumefaciens* strain GV3101 (pMP90) that were kept on ice. The solution was transferred to an electroporation cuvette (on ice). The cells were electroporated at 2.40 kV for 9.70 msec. Then, approximately 750 μl of SOC media was added to the cells. The cells were resuspended and transferred to culture tubes where they were incubated at 28°C for 2 hours. The cells were then spread onto 2% LB agar plates containing the antibiotics listed below.

LB media for *A. tumefaciens* transformation was prepared as follows: 1% tryptone, 0.5% yeast extract, 0.5% sodium chloride, buffered with 10 mM MES and adjusted to pH 5.7. Gentamicin (for the *A. tumefaciens* helper plasmid), rifampicin (for *A. tumefaciens*), and either kanamycin (for cells transformed with pCAMBIA) or spectinomycin (for cells transformed with pSPDK) were added at concentrations of 100 μg/ml, 50 μg/ml, 50 μg/ml, and 50 μg/ml, respectively.

Plant growth conditions

Plants were grown on soil or on 0.8% MS agar plates with 1.0% sucrose, as described below. Conditions in the growth room were as follows: 18h light/6h darkness (approximately 100 μmol · m^{-2} · s^{-1}), a temperature of 23°C during the light phase and 18°C during the dark phase, and 70% humidity. Wild type plants were of the Columbia ecotype.

Transformation of *Arabidopsis* plants with *A. tumefaciens*

        A. tumefaciens was cultured in LB media (as described above) overnight. The cell culture was diluted to an OD_{600} of 0.8-1.0 and added to plant transformation mix (2.15 g/L MS salt, 5% sucrose, 0.03% Silwet L-77). Plant inflorescences were dipped into the transformation mix for 1-2 minutes, then removed and laid horizontally in trays lined with moist paper towels. The trays were covered with plastic wrap. The plants were left in the growth chamber overnight. The plastic wrap was removed and the plants were returned to a vertical position.

Selection of T1 seeds from T0 plants

        Approximately 200 mg of T1 seeds from transformed T0 generation plants were collected and placed in 50 ml Falcon tubes. 10 ml of 75% ethanol was added to the tubes and decanted after 1 minute. Then 10 ml of 30% bleach and 0.05% Tween 20 were added to the tubes. The tubes were placed on a shaker for 15 minutes, during which time the contents of individual tubes were periodically mixed by manually rotating them for 45-60 seconds at a time. Afterwards, the liquid from the tubes was decanted. The seeds were washed 6-8 times with approximately 30 ml of autoclaved, double distilled water. The seeds were then resuspended in roughly 25 ml of 0.2% MS agar that had been autoclaved at 121°C and maintained at 65°C. The resuspended seeds were spread onto 0.8% MS agar plates containing the ingredients listed below. The plates were sealed with Parafilm and transferred to a growth room. Care was taken to ensure a sterile environment for the plates. Seedlings were allowed to sprout for 3-4 weeks before being transplanted to soil.

The 0.8% MS agar plates were prepared as follows: 0.43% MS salt, 1.0% sucrose, 0.05% MES, pH 5.7 (adjusted using KOH), and 0.8% agar. After autoclaving, Gamborg's vitamin solution was added to a final concentration of 1 ml/L. In addition, vancomycin and either hygromycin (for plants transformed with pCAMBIA vector) or gentamicin (for plants transformed with pSPDK vector) were added at final concentrations of 500 μg/ml, 50 μg/ml and 100 μg/ml, respectively, to select for transformed seedlings. The plates were left overnight at 4°C before use.

Ethanol extraction of leaf samples for carbohydrate analysis

For each sample, 2-3 leaves were collected and placed in 1.5 ml microfuge tubes. If storage was necessary at this point, the samples were placed in dry ice at -80°C. Otherwise, preparation of the samples continued as follows. Approximately 1 ml of 70% ethanol was added to the tubes. The tubes were incubated at 70°C for 30 minutes, then the supernatant was decanted. These steps were repeated once (approximately 1 ml of 70% ethanol for 30 minutes). The tubes were left open and allowed to air dry overnight. The samples were stored at room temperature.

Fractionation of leaf material

This method is based on the procedure described by Reiter, et al. (1997). Approximately 10 g of leaf material was collected in a Magenta box with liquid nitrogen, then stored at -80°C.

Before beginning the fractionation, a mortar and pestle were stored at -80°C overnight. Upon removal, they were pre-cooled with liquid nitrogen. The leaf material

was placed in the mortar. Then liquid nitrogen was added and the mixture was ground. This was repeated 2-3 times. The powderized mixture was placed in a Magenta box and stored at -80°C. A small amount was resuspended in water and viewed under a microscope to ensure that the leaf material had been sufficiently homogenized.

The pulverized leaf material was resuspended in 50 ml of buffer A (100 mM MOPS, pH 7.0, 1.5% SDS, 5 mM sodium bisulfite that was heated briefly at approximately 50°C to remove turbidity), and stirred in an ice-water bath. The cell wall material was pelleted by centrifugation for 10 minutes at 4°C in a Sorvall SS34 rotor at 15000 rpm. The material was then washed with 25 ml of ice-cold buffer B (100 mM MOPS, pH 7.0, 0.5% SDS, 3 mM sodium bisulfite), and centrifuged as described above. The two supernatants were combined into fraction S1.

The cell wall pellet was washed with 50 ml of buffer B, and the supernatant was discarded after centrifugation. Following 2-3 washes with water, the cell wall material was resuspended in 16 ml of water and stored at -20°C.

A solution of 10 M ammonium acetate was added to fraction S1 to have a final concentration of 1 M. Polymers were precipitated by addition of five volumes of a 1:1 mixture of ethanol and acetone. The mixture was allowed to stand for 24 hours at room temperature. The precipitate was then recovered by centrifugation at 2000 g for 15 minutes. Neither the addition of more acetone nor the addition of more ammonium acetate to the mixture caused turbidity, suggesting that polymers had been efficiently removed. The pellet was washed thoroughly with 80% ethanol, and stored at 4°C as a suspension in 40 ml of 80% ethanol.

Samples were prepared for carbohydrate analysis by properly resuspending them and pipetting an amount of material roughly equal to 1% (approximately 100 mg) of the original mass of the leaves into microfuge tubes. The samples were centrifuged to pellet the insoluble material. The supernatant was removed and the remaining material was dried by centrifugation under vacuum in a SpeedVac apparatus. The samples were stored at room temperature.

<u>Cell wall carbohydrate analysis</u>

This method is based on the cell wall analysis method described by Blakeney, et al. (1983).

Each sample of cell wall material was autoclaved for one hour in 250 μl of 2 M TFA with slow exhaust. To each sample, 100 μg of *myo*-inositol were added for use as an internal standard. For calibration purposes, a mixture of monosaccharides (100 μg each of L-rhamnose, L-fucose, L-arabinose, D-xylose, D-mannose, D-galactose and D-glucose) with 100 μg *myo*-inositol in 250 μl of 2 M TFA was used. In addition, blank samples (2 M TFA with 100 μg *myo*-inositol) were autoclaved.

Samples were stored overnight at 4°C and then transferred to large screw cap vials with Teflon-lined lids. 100 μl of 9 M NH_3 was added to each tube and briefly vortexed. The tubes were transferred to a 40°C water bath. 1 ml of 2% $NaBH_4$ in DMSO was added to each tube, followed by brief vortexing. The tubes were then covered with aluminum foil, and allowed to incubate at 40°C for 90 minutes.

The tubes were removed from the water bath. Following the addition of 250 μl of acetic acid, the tubes were briefly vortexed. 250 μl of 1-methylimidazole and 4 ml of

acetic anhydride were added to each tube. The tubes were vortexed and incubated at room temperature for 10 minutes. Eight milliliters of double distilled water were added to each tube. The tubes were capped, inverted several times to dissolve a precipitate, and allowed to cool slowly to below 30°C. 1.5 ml CH_2Cl_2 was added to each tube, and the alditol acetates were extracted into the organic phase by inverting the tubes at least 30 times. The tubes were allowed to sit at 4°C for 2 hours (or overnight).

The lower (CH_2Cl_2) phases were recovered using long Pasteur pipets and transferred to small (12 x 75 mm) test tubes. 200 μl of CH_2Cl_2 followed by 2 ml of double distilled water were added to each tube. The tubes were vortexed vigorously for over 15 seconds. After the organic phase settled it was transferred, via a P-200 pipettor, to small pre-numbered GC vials. The filled vials were stored at -20°C.

The samples were analyzed by gas-liquid chromatography essentially as described by Reiter, et al. (1997) with minor modifications. The inlet temperature for the analysis was set to 240°C, the detector temperature was set to 250°C, and the total run time was 22 minutes.

Subcellular localization of GT14

The amino acid sequences of GT14 and MUR3 were analyzed by the TMHMM

online tool at http://www.cbs.dtu.dk/services/TMHMM/ (Fig. 5). The algorithm uses

bioinformatics data to predict the hydrophobicity of each amino acid in the sequence.
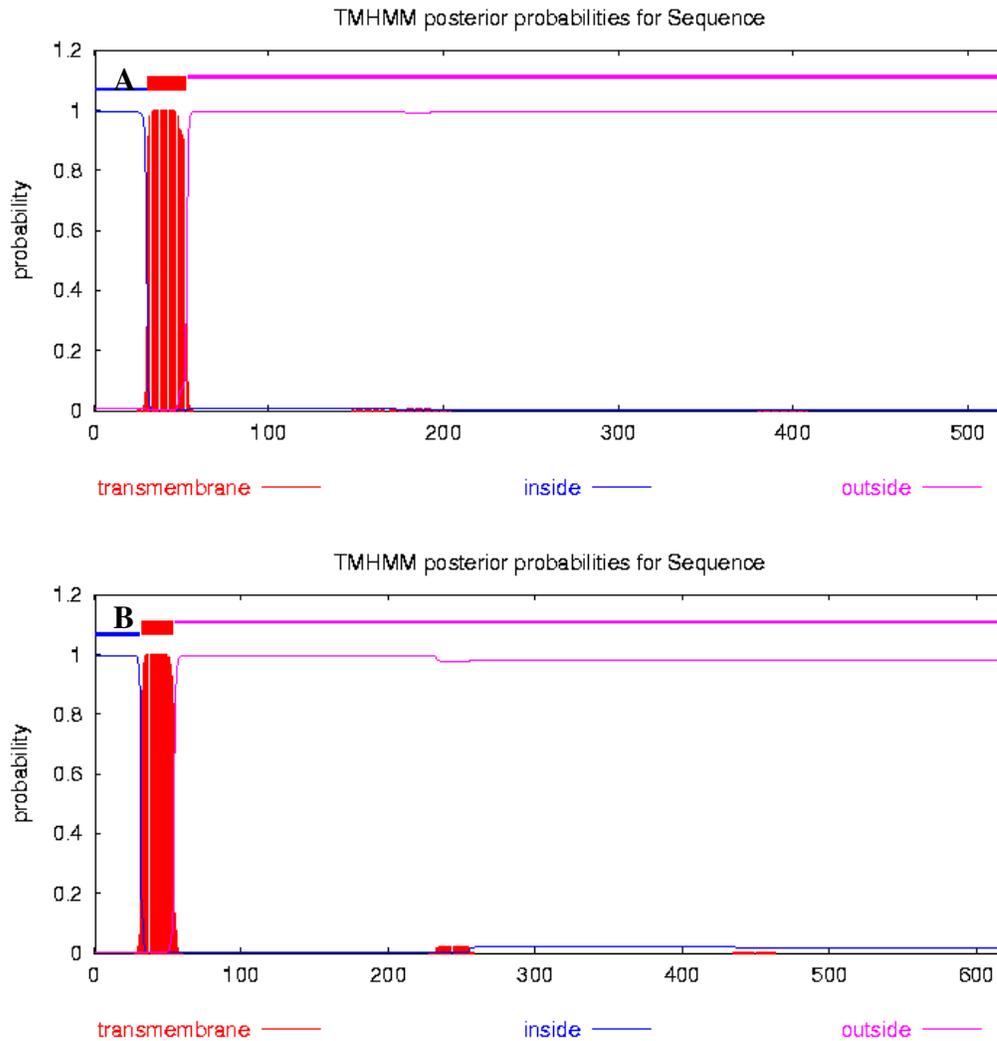
Figure 5. The amino acid sequences of GT14 **(A)** and MUR3 **(B)** were analyzed by the TMHMM transmembrane prediction algorithm. The resulting plots predict that amino acids 31-53, in GT14, and amino acids 32-54, in MUR3, comprise transmembrane domains.

Hydrophobic regions of peptides are generally embedded in the lipid bilayer. The

algorithm determines the likelihood of any given amino acid in the sequence being part of

such an α-helical transmembrane domain. The tool predicted that, in GT14, amino acids

31 through 53 comprise an α-helical transmembrane domain (Fig 5A). In MUR3, amino

acids 32 through 54 were predicted to be part of an α-helical transmembrane domain (Fig.

5B).

The amino acid sequences of GT14 and MUR3 were also analyzed by the SignalP

3.0 online tool (Fig. 6) at http://www.cbs.dtu.dk/services/SignalP/ (Bendtsen et al., 2004;

Nielsen et al., 1997).
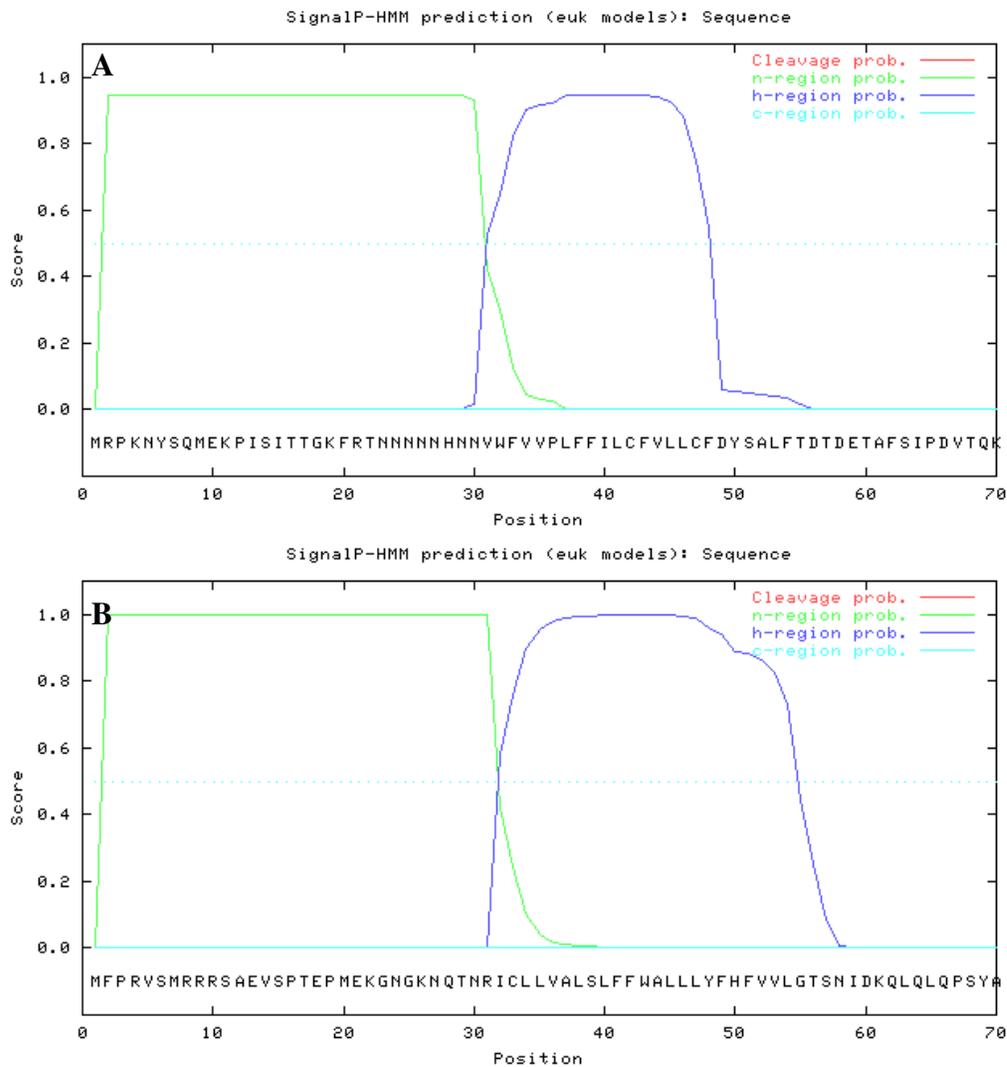
Figure 6. The amino acid sequences of GT14 and MUR3 were analyzed by the SignalP 3.0 algorithm. The resulting plot predicts the presence of a signal anchor sequence (*p=0.947*) for GT14 **(A)** as well as the presence of a signal anchor sequence (*p=0.999*) for MUR3 **(B)**.

The SignalP hidden Markov model (Nielsen and Krogh, 1998) output predicts with a probability of 0.947 that the polypeptide contains a signal anchor sequence. This is based on the presence of an h-region (hydrophobic region), roughly between amino acids 30 and 50, preceded by an n-region. For MUR3, it predicts the presence of a signal anchor with probability 0.999. This is indicated by an h-region, roughly between amino acids 30 and 60, preceded by an n-region. In both cases, the n-region and subsequent h-region suggest the presence of a sequence responsible for targeting the peptide to a cellular membrane. These data correspond to the output obtained from the TMHMM tool. It should be noted that MUR3 is known to be targeted to the Golgi membrane (Tamura et al., 2005).

Carbohydrate composition of gt14 mutants

The relative amounts of monosaccharides in leaf material from *gt14* and WT

plants were quantified via gas-liquid chromatography (Fig. 7).
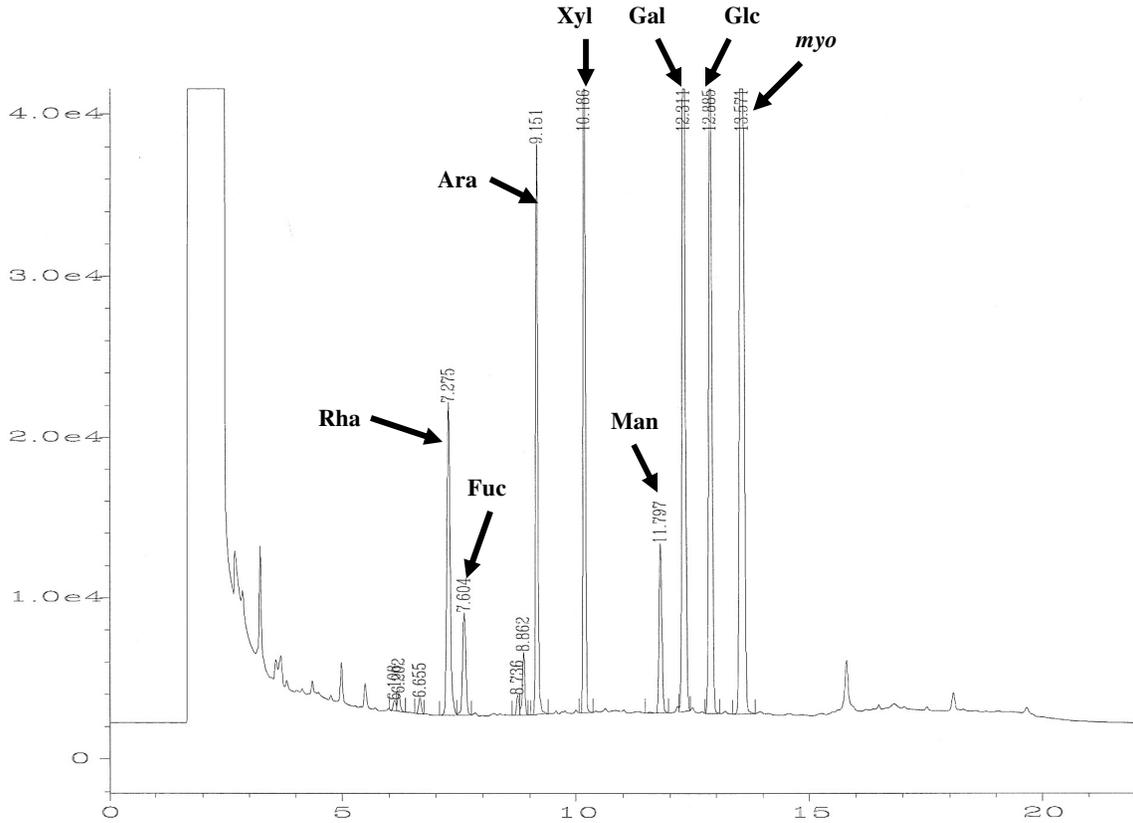
Figure 7. Representative image of a chromatogram obtained by gas-liquid chromatography of the derivatized monosaccharides. Rhamnose, fucose, arabinose, xylose, mannose, galactose, glucose and *myo*-inositol were separated based on differences in their retention times, and quantified based on their respective peak areas.

Total cell wall material from *gt14* plants displayed 20.7% and 12.8% decreases in

fucose and rhamnose, respectively, as well as 30.6% and 12.5% increases in mannose and

arabinose, respectively (Fig. 8 and Table 1). There was no significant difference between

the *gt14* mutants and WT plants with regard to xylose or galactose.
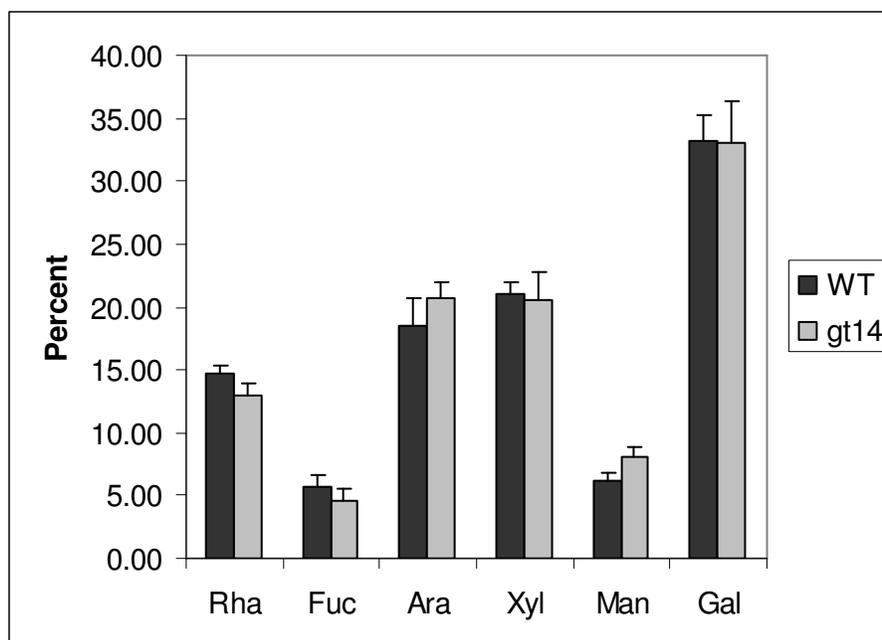
Figure 8. Cell wall monosaccharide composition in *gt14* and WT plants. The *gt14* mutants displayed significant decreases in fucose and rhamnose, as well as significant increases in mannose and arabinose.

Monosaccharide	*gt14*	Wild type	Change
Rhamnose	12.9%	14.8%	– 12.8%
Fucose	4.6%	5.8%	– 20.7%
Arabinose	20.7%	18.4%	+ 12.5%
Xylose	20.6%	21.0%	– 1.9%
Mannose	8.1%	6.2%	+ 30.6%
Galactose	33.0%	33.2%	– 0.6%

Table 1. Relative amounts of monosaccharides in *gt14* and WT plants.

Differences between the mutant and WT plants were also evaluated by constructing histograms of the data for each monosaccharide (Fig. 9). In the *gt14* plants, there was a distinct decrease in the median amount of fucose (4.5% vs. 6%). There was also an increase in the median amount of mannose (8.5% vs. 6%), a decrease in the median amount of rhamnose (12% vs. 15%), and an increase in the median amount of arabinose (20% vs. 18%). There was, however, a relatively large amount of overlap between the WT and *gt14* distributions with regard to rhamnose and arabinose, so the differences are less clear than they may initially seem.
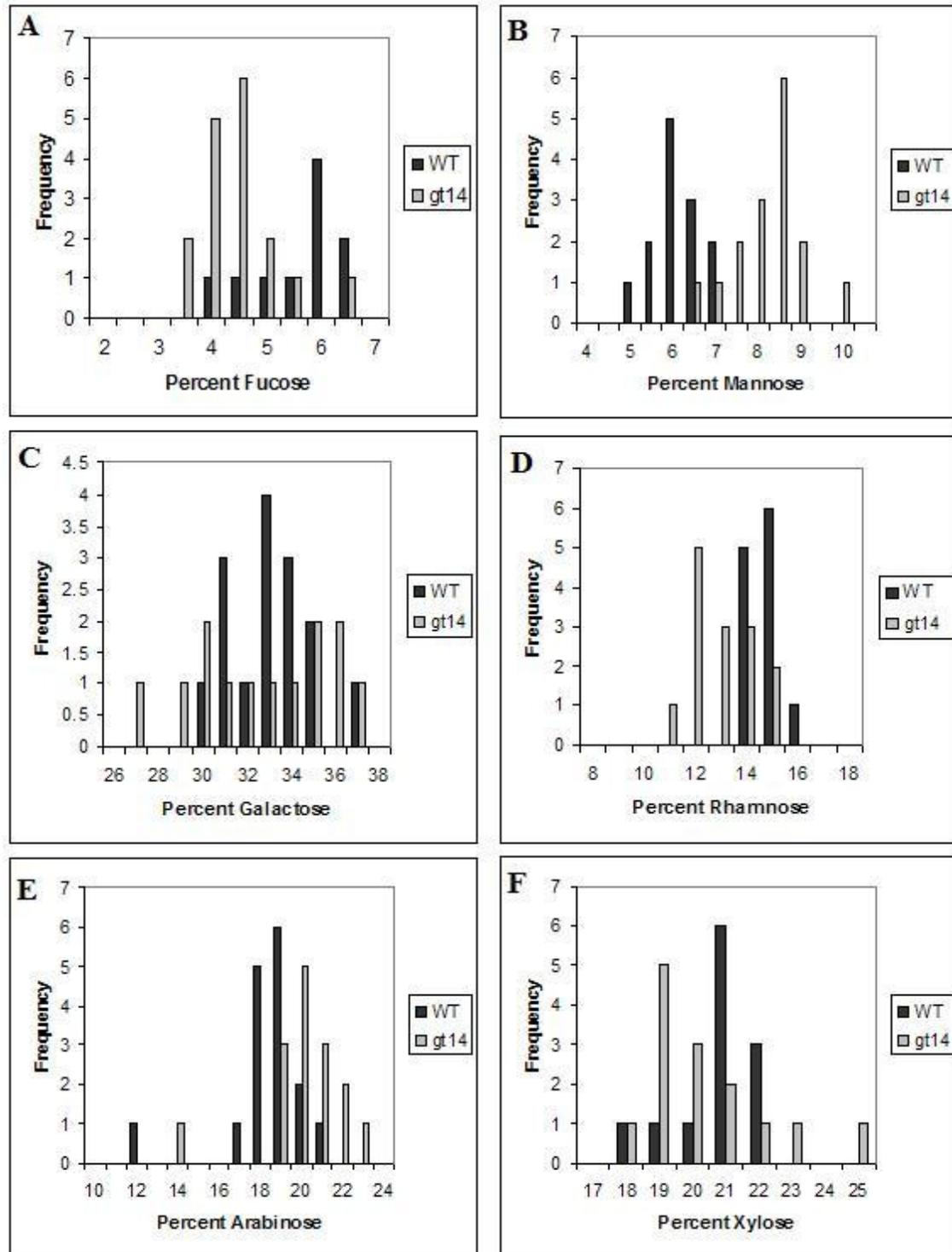
Figure 9. Histograms of the carbohydrate analysis of *gt14* and WT plants. There were relatively large shifts in the median for fucose, which decreased, and mannose, which increased. There were slightly smaller shifts in the median for rhamnose, which decreased, and arabinose, which increased.

28

In addition, a difference is noticeable in xylose content (19% vs. 21%). It can be observed, however, that these distributions are skewed with regard to xylose (*gt14* to the right and WT to the left) which, perhaps, makes these differences less significant. There was no observable change in the distribution of galactose between the two groups.

In order to determine whether or not the changes in monosaccharide composition were the result of changes in the cell wall or in the cytosolic material, fractionated portions of cell wall material and soluble polymers were analyzed using gas-liquid chromatography (Fig. 10).

In the *gt14* plants, less fucose (Fig. 10A) was present in the cell wall than in the WT (ca. 4% vs. 4.5%) but in relative terms, there was more fucose in the *gt14* soluble fraction than in the wild type soluble fraction (ca. 7% vs. 5.5%). The amount of mannose in the *gt14* and WT cell wall fractions was approximately the same (Fig. 10B), though there was significantly more in the *gt14* soluble fraction than in the WT soluble fraction (ca. 9% vs. 6%). The *gt14* cell wall fraction contained more galactose (Fig. 10C) than the WT cell wall fraction (ca. 33% vs. 26%), though the amount of galactose in the soluble fraction was similar in both sets of plants. There was slightly less rhamnose (Fig. 10D) in both fractions for *gt14*. There was only slightly more arabinose (Fig. 10E) in the *gt14* cell wall fraction than in the WT cell wall fraction, and slightly more arabinose in the WT soluble fraction than in the *gt14* soluble fraction. There was no significant change in the xylose content (Fig. 10F) of the soluble fraction, but there was slightly more xylose in the WT cell wall fraction than in the *gt14* cell wall fraction.

It should be noted that sample sizes in the fractionation analysis were small. This may explain some peculiarities in the histograms, such as the lack of a distinct median or

peak in the *gt14* distribution of cell wall rhamnose (Fig. 10D). Thus, certain changes may

not necessarily be significant.

gt14 cell wall
WT cell wall
gt14 soluble polymers
WT soluble polymers
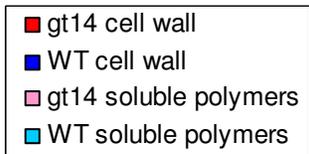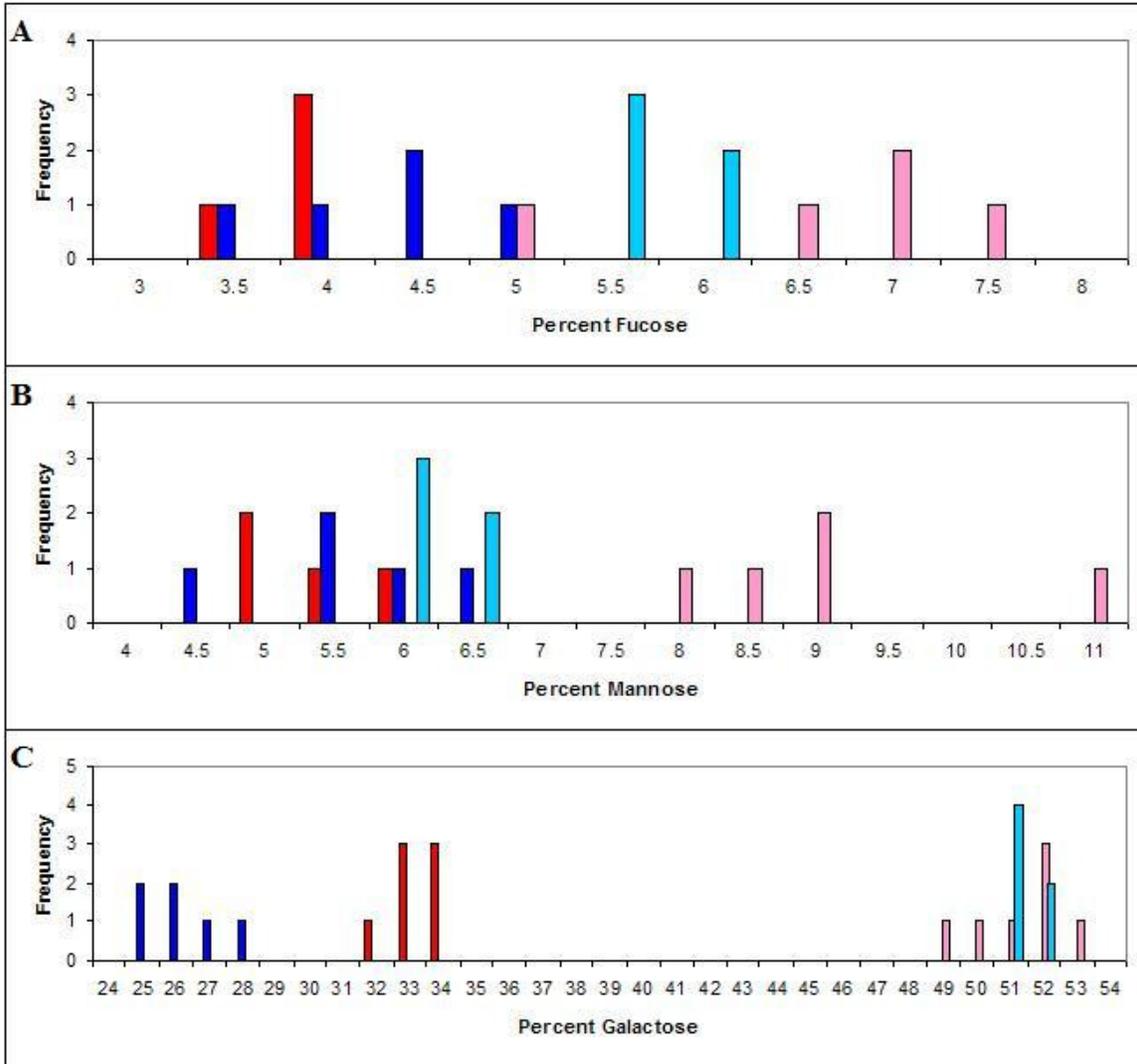
Figure 10. Leaf material from *gt14* and WT plants was fractionated to separate cell wall components from soluble polymers. **(A)-(F)** present histograms of the distribution of monosaccharides between the two fractions.

Figure 10. (Continued)

- ■ gt14 cell wall
- ■ WT cell wall
- ■ gt14 soluble polymers
- ■ WT soluble polymers
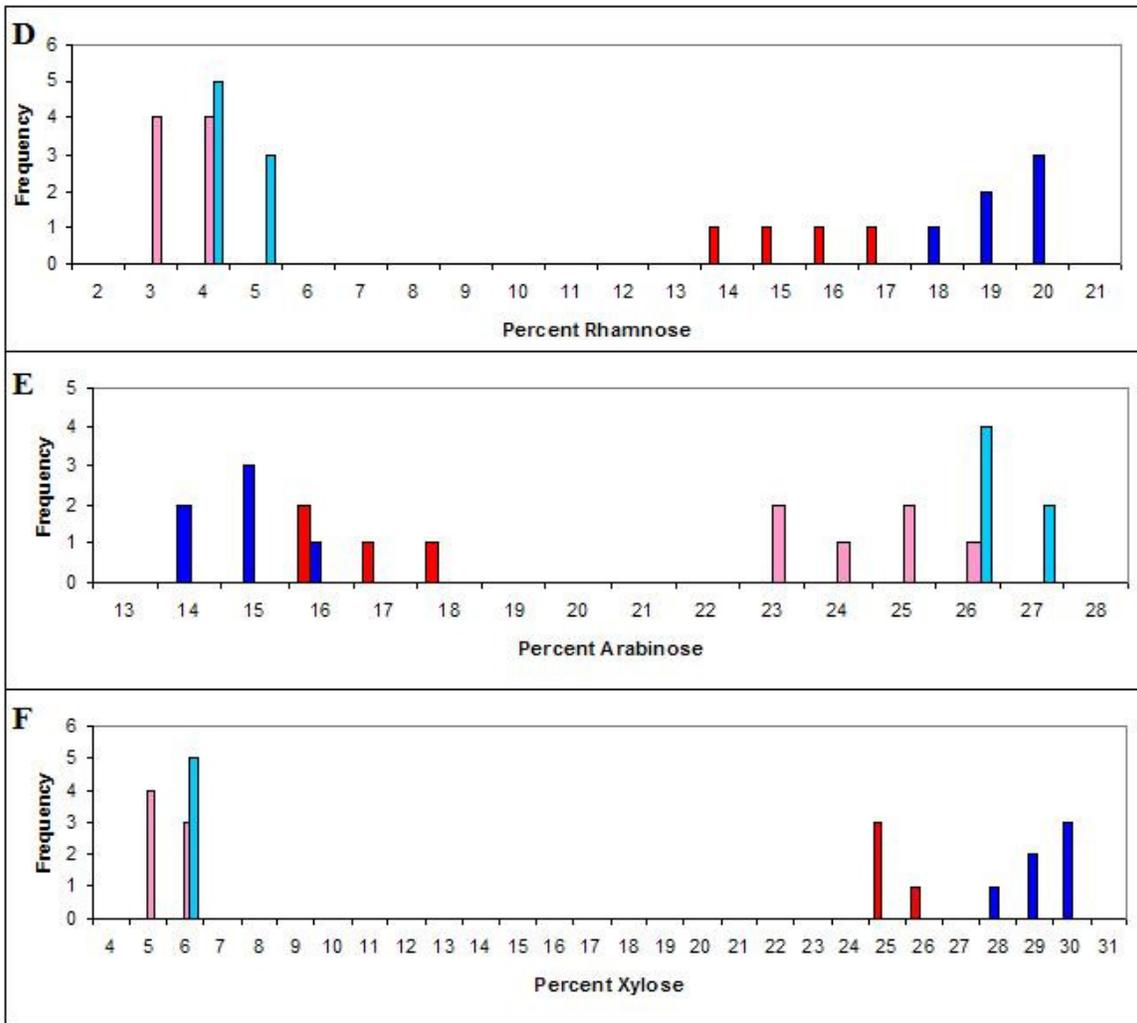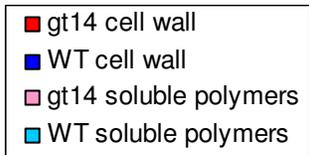
Complementation and overexpression analysis

The *GT14* coding region was inserted into pCAMBIA 1301 and pSPDK vectors and then transformed into WT and *gt14* plants. T1 seeds were collected from T0 generation plants. The monosaccharide composition of the T1 transformants was analyzed via gas-liquid chromatography of alditol acetates. Three sets of plants were analyzed: *gt14* plants transformed with the *GT14*-pSPDK construct (*gt14/GT14*-pSPDK), WT plants transformed with the *GT14*-pSPDK (WT/*GT14*-pSPDK), and *gt14* plants transformed with the *GT14*-pCAMBIA construct (*gt14/GT14*-pCAMBIA), in addition to their respective controls. A set of WT plants transformed with the *GT14*-pCAMBIA construct is not included because leaf material for these plants was not ready to be harvested at the time of this writing. Histograms of the data (Figs. 13-18) are located in the Appendix.

There was relatively less fucose in the *gt14/GT14*-pSPDK plants than in the control group (Fig. 13A). The same relationship was observed in the WT/*GT14*-pSPDK plants (Fig. 13B). There was no change in the distribution for fucose in the *gt14/GT14*-pCAMBIA groups (Fig. 13C). The *gt14/GT14*-pSPDK plants (Fig. 14A) actually appeared to have somewhat more mannose than the control, as did the WT/*GT14*-pSPDK plants (Fig. 14B). There was no difference in the relative amount of mannose between the *gt14/GT14*-pCAMBIA and control groups (Fig. 14C). The distributions of galactose in the *gt14/GT14*-pSPDK (Fig. 15A) and *gt14/GT14*-pCAMBIA (Fig. 15C) groups revealed no changes either, though the amount of galactose was noticeably lower (approximately 30% vs. 35%) in the WT/*GT14*-pSPDK group (Fig. 15B). None of the transformants

32

showed considerable differences in rhamnose (Fig. 16), arabinose (Fig. 17) or xylose (Fig. 18).

The age of the plants at the time of leaf material collection was not uniform. Some transformants were older than others. This, combined with small sample sizes, may have contributed to the variability between the samples and the skewed appearance of the distribution data.

Discussion

GT14 appears to be a type II membrane protein most likely targeted to the Golgi apparatus. This cannot be ascertained based on the bioinformatics data alone; however, it is probable because the TMHMM and SignalP results for GT14 (Fig. 5A and Fig. 6A) are nearly identical to those for MUR3 (Fig. 5B and Fig. 6B). MUR3 is known to be a type II membrane protein with a short cytosolic domain and a large lumenal domain. It is targeted to the Golgi apparatus (Tamura et al., 2005). It would not be unreasonable to assume that GT14 occupies a similar position in the Golgi membrane. The *gt14* mutants showed a reduction in fucose of approximately 20.7% and an increase in mannose of approximately 30.6%. Analysis of cell wall and soluble fractions confirmed that the reduction in fucose could be traced to the cell wall. Considering these changes within the context of phylogenetic relationships may provide clues as to the exact function of GT14. *GT14*, *MUR3*, and *GT18* are all part of subgroup A in the GT47 gene family (Fig. 11).
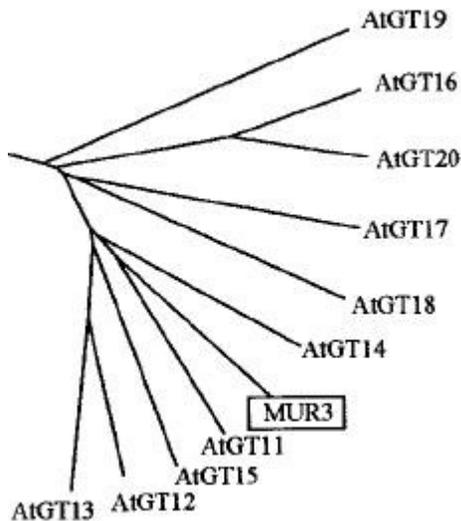
Figure 11. Subgroup A of the GT47 family. Subgroup A includes *GT14* in addition to its homologs *MUR3* and *GT18*, among others (Li et al., 2004).

MUR3 is a xyloglucan specific galactosyltransferase and, based on genetic data, GT18 is as well. Therefore, it is relatively safe to assume that the target of GT14 is also

xyloglucan. The exact nature of its function, though, remains unclear. If GT14 is a galactosyltransferase like MUR3 and GT18, the decrease in fucose in *gt14* mutants may be explained as an indirect consequence of such activity. Since galactosylated xyloglucan is an acceptor substrate for the fucosyltransferase AtFUT1, a decrease in the galactosylation of xyloglucan is also manifested as a reduction in fucosylated xyloglucan. This effect is seen clearly in the *mur3* mutants. Although MUR3 is a galactosyltransferase, the most pronounced change in *mur3* plants is a 50% reduction in fucose content (Madson et al., 2003). The same may be true of GT14. Unlike the *mur3* and *gt18* mutants, the *gt14* mutants did not display a reduction in cell wall galactose. This does not rule out the potential role of GT14 as a xyloglucan-specific galactosyltransferase. It is entirely possible that *gt14* mutants contain less galactosylated xyloglucan, but that this decrease is masked by galactose in other cell wall components like RG-I and arabinogalactans.

One way to gain insight into the function of GT14 may be to observe the structural and mechanical characteristics of *gt14* mutants, and then compare these to the characteristics of *mur3* mutants. For instance, growing *mur*3 mutants were found to be visually indistinguishable from their wild type counterparts, save for a collapsed appearance of the trichome papillae (Madson et al., 2003). Moreover, there was no difference in the amount of force needed to break elongating inflorescences in *mur3* mutants (Madson et al., 2003). While the *mur3* mutation does not appear to be detrimental in relatively older plants, there was, on the other hand, a considerable decrease in the tensile strength of the hypocotyls in 4 day old *mur3* seedlings (Peña et al., 2004). Given these qualities of *mur3* mutants, studies on the trichome structure and

mechanical strength of *gt14* mutants at different stages of development would be beneficial. If the *gt14* mutants and *mur3* mutants display similar trends, then it would strengthen the hypothesis that GT14 and MUR3 possess a similar function.

Since mutations in both *GT14* and *MUR3* reduce fucose content, it is possible that GT14 and MUR3 (or other proteins in subgroup A) cooperate in the galactosylation of xyloglucan, though it is also possible that they function independently of one another, to some extent. Further experimentation is necessary to determine if there are any significant interactions between the two. It may be useful to develop and analyze *mur3 gt14* double mutants for this purpose.

The increase in the mannose content in the soluble polymer fraction of the *gt14* mutants was significant as well. This may potentially be explained by the fact that D-mannose is a substrate in the synthesis of L-fucose (Fig. 12). GDP-D-mannose is converted to GDP-L-fucose by means of GDP-D-mannose 4,6-dehydratase (Mulichak et al., 2002; Bonin et al., 2003). It is unlikely that GT14 is involved in this process; however, *gt14* plants demonstrate a marked reduction in fucose. If less GDP- L-fucose is incorporated into xyloglucan in the cell wall, this may mean that, to compensate, less GDP-D-mannose is converted to GDP-L-fucose. This could lead to an increased concentration of GDP-D-mannose within the cell, though the precise nature of this effect would be unclear. Two ways this may be manifested include reductions in the transcription and translation of the *GMD2* gene, which encodes for GDP-D-mannose 4,6-dehydratase. These reductions could be measured by assays that detect changes in transcription of *GMD2* into mRNA (e.g., Northern blot) and translation of the mRNA into GDP-D-mannose 4,6-dehydratase (e.g., Western blot).
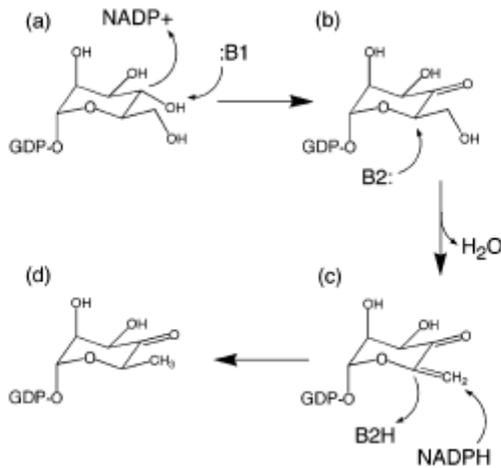
Figure 12. A proposed mechanism for the catalytic activity of GDP-D-mannose 4,6-dehydratase. GDP-mannose (a) is converted to GDP-4-keto-6-deoxymannose (d) via the 4-keto and 4-keto-5,6-ene (b, c) intermediates (Mulichak et al., 2002).

Complementation and overexpression analysis was ultimately inconclusive. A number of factors likely contributed to this. The T1 plants from which the leaf material was harvested were older than the *gt14* plants that were originally used in the determination of carbohydrate composition. Fucose and xylose content have been shown to increase with the age of the plant, whereas arabinose and galactose appear to decrease with age (Reiter et al., 1997). This effect may have dampened some of the differences in monosaccharide composition and exaggerated others. In addition, since T1 generation plants were used, the majority of plants were most likely hemizygous for the *GT14-* plasmid constructs, rather than homozygous. T2 generation plants, which should be homozygous, may display more noticeable phenotypic differences, if any.

Further studies are necessary for a better understanding of GT14 and its activity. The composition of different types of xyloglucan in the mutants remains be analyzed. Knowledge of the relative amounts of the different forms of xyloglucan (XXXG, XLXG, XXLG, XLLG, XXFG, XLFG) in the *gt14* mutants may shed much light on the function of GT14. *In vitro* enzyme assays of GT14 on xyloglucan polysaccharide may also

indicate more precisely which step in the galactosylation of xyloglucan GT14 is involved in. GUS-tagged (as in the pCAMBIA construct) or citrine-tagged (as in the pSPDK construct) *GT14* may also be used to ascertain whether or not GT14 is targeted to the Golgi membrane. This could be accomplished by fractionating the cellular organelles, and then using an immunohistochemical assay like Western blotting or ELISA to determine if the tagged protein is present in a particular fraction. GT14 may also possess other functions. For example, it is known that a different defect in the *MUR3* gene, *kam1*, affects cytoskeletal organization (Tamura et al., 2005). In addition, *mur3* mutants demonstrated enhanced resistance to *H. parasitica* in petiole tissue. The mechanism for this pathogenic resistance is not entirely clear. One hypothesis is that proteins for defense against pathogens are activated by perturbations in the cell wall or in the actin cytoskeleton and that the loss of MUR3 conditions these proteins, essentially lowering the threshold at which they are activated. Another alternative is that there is some molecule that activates pathogen defenses, known as an 'elicitor,' normally produced by pathogen activity, and that this elicitor is also produced by changes to the cell wall (Tedman-Jones et al., 2008). Though the mechanisms for these hypotheses are unclear, the homology between MUR3 and GT14 suggests that GT14 may also be involved in such processes.

Ultimately, this work has revealed several characteristic features of GT14. Still, many questions regarding its mechanism and function remain to be answered.

References

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucl. Acids Res. 25: 3389-3402.

Bacic A, Harris PJ, Stone BA. 1988. Structure and function of plant cell walls. *In* PK Stumpf, EE Conn, eds, The Biochemistry of Plants, Vol 14. Academic Press, New York, pp 297-371.

Bendtsen JD, Nielsen H, von Heijne G, Brunak S. 2004. Improved prediction of signal peptides: SignalP 3.0. J. Mol. Biol. 340: 783-795.

Blakeney AB, Harris PJ, Henry RJ, Stone BA. 1983. A simple and rapid preparation of alditol acetates for monosaccharide analysis. Carbohydr. Res. 113: 291-299.

Bonin CP, Freshour G, Hahn MG, Vanzin GF, Reiter WD. 2003. The GMD1 and GMD2 genes of Arabidopsis encode isoforms of GDP-D-mannose 4,6-dehydratase with cell type-specific expression patterns. Plant Physiol. 132: 883-892.

Brown DM, Zeef LA, Ellis J, Goodacre R, Turner SR. 2005. Identification of novel genes in Arabidopsis involved in secondary cell wall formation using expression profiling and reverse genetics. Plant Cell 17: 2281-2295.

Carpita N, Tierney M, Campbell M. 2001. Molecular biology of the plant cell wall: searching for the genes that define structure, architecture and dynamics. Plant Mol. Biol. 47: 1–5.

Carpita NC, Gibeaut DM. 1993. Structural models of primary cell walls in flowering plants: consistency of molecular structure with the physical properties of the walls during growth. Plant J. 3: 1-30.

Delmer DP. 1999. Cellulose biosynthesis: exciting times for a difficult field of study. Annu. Rev. Plant Physiol. Plant Mol. Biol. 50: 245-276.

Freshour G, Bonin CP, Reiter WD, Albersheim P, Darvill AG, Hahn MG. 2003. Distribution of fucose-containing xyloglucans in cell walls of the *mur1* mutant of Arabidopsis. Plant Physiol. 131: 1602-1612.

Fry SC. 1994. Oligosaccharins as plant growth regulators. Biochem. Soc. Symp. 60: 5-14.

Kobayashi M, Matoh T, Azuma J. 1996. Two chains of rhamnogalacturonan II are cross-linked by borate-diol ester bonds in higher plant cell walls. Plant Physiol. 110: 1017-1020.

Levy S, Maclachlan G, Staehelin LA. 1997. Xyloglucan sidechains modulate binding to cellulose during in vitro binding assays as predicted by conformational dynamics simulations. Plant J. 11: 373-386.

Levy S, York WS, Stuike-Prill R, Meyer B, Staehelin LA. 1991. Simulations of the static and dynamic molecular conformations of xyloglucan. The role of the fucosylated side-chain in surface-specific side-chain folding. Plant J. 1: 195-215.

Li X, Cordero I, Caplan J, Mølhøj M, Reiter WD. 2004. Molecular analysis of 10 coding regions from Arabidopsis that are homologous to the MUR3 xyloglucan galactosyltransferase. Plant Physiol. 134: 940-950.

Madson M, Dunand C, Li X, Verma R, Vanzin GF, Caplan J, Shoue DA, Carpita NC, Reiter WD. 2003. The *MUR3* gene of Arabidopsis encodes a xyloglucan galactosyltransferase that is evolutionarily related to animal exostosins. Plant Cell 15: 1662-1670.

McCann MC, Roberts K. 1991. Architecture of the primary cell wall. *In* CW Lloyd, ed, The Cytoskeletal Basis of Plant Growth and Form. Academic Press, London, pp 109-129.

McNeil M, Darvill A, Fry SC, Albersheim P. 1984. Structure and function of the primary cell wall of plants. Annu. Rev. Biochem. 53: 625-663.

Mulichak AM, Bonin CP, Reiter WD, Garavito RM. 2002. Structure of the MUR1 GDP-mannose 4,6-dehydratase from Arabidopsis thaliana: implications for ligand binding and specificity. Biochemistry 41: 15578-15589.

Nielsen H, Krogh A. 1998. Prediction of signal peptides and signal anchors by a hidden Markov model. Proc. Int. Conf. Intell. Syst. Mol. Biol. (ISMB 6), AAAI Press, Menlo Park, California pp. 122-130

Nielsen H, Engelbrecht J, Brunak S, von Heijne G. 1997. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Eng. 10: 1-6.

O'Neill MA, Eberhard S, Albersheim P, Darvill AG. 2001. Requirement of borate cross-linking of cell wall rhamnogalacturonan II for Arabidopsis growth. Science 294: 846-849.

Pauly M, Eberhard S, Albersheim P, Darvill A, York WS. 2001. Effects of the mur1 mutation on xyloglucans produced by suspension-cultured Arabidopsis thaliana cells. Planta 214: 67-74.

Peña MJ, Ryden P, Madson M, Smith AC, Carpita NC. 2004. The galactose residues of xyloglucan are essential to maintain mechanical strength of the primary cell walls in Arabidopsis during growth. Plant Physiol. 134: 443-451.

Reiter WD. 2002. Biosynthesis and properties of the plant cell wall. Curr. Opin. Plant Biol. 5: 536-542.

Reiter WD, Chapple C, Somerville CR. 1997. Mutants of Arabidopsis thaliana with altered cell wall polysaccharide composition. Plant J. 12: 335-345.

Reiter WD, Chapple C, Somerville CR. 1993. Altered growth and cell walls in a fucose-deficient mutant of Arabidopsis. Science 261: 1032-1035.

Ryden P, Sugimoto-Shirasu K, Smith AC, Findlay K, Reiter WD, McCann MC. 2003. Tensile properties of Arabidopsis cell walls depend on both a xyloglucan cross-linked microfibrillar network and rhamnogalacturonan II-borate complexes. Plant Physiol. 132: 1033-1040.

Tamura K, Shimada T, Kondo M, Nishimura M, Hara-Nishimura I. 2005. KATAMARI1/MURUS3 is a novel golgi membrane protein that is required for endomembrane organization in Arabidopsis. Plant Cell 17: 1764-1776.

Tedman-Jones JD, Lei R, Jay F, Fabro G, Li X, Reiter WD, Brearley C, Jones JD. 2008. Characterization of Arabidopsis mur3 mutations that result in constitutive activation of defence in petioles, but not leaves. Plant J. 56: 691-703.

Turner SR, Somerville CR. 1997. Collapsed xylem phenotype of Arabidopsis identifies mutants deficient in cellulose deposition in the secondary cell wall. Plant Cell 9: 689–701.

Vanzin GF, Madson M, Carpita NC, Raikhel NV, Keegstra K, Reiter WD. 2002. The *mur2* mutant of *Arabidopsis thaliana* lacks fucosylated xyloglucan because of a lesion in fucosyltransferase AtFUT1. Proc. Natl. Acad. Sci. USA 99: 3340-3345.

York WS, Darvill AG, Albersheim P. 1984. Inhibition of 2,4-dichlorophenoxyacetic acid-stimulated elongation of pea stem segments by a xyloglucan oligosaccharide. Plant Physiol. 75: 295-297.

Zablackis E, York WS, Pauly M, Hantus S, Reiter WD, Chapple CCS, Albersheim P, Darvill A. 1996. Substitution of L-fucose by L-galactose in cell walls of *Arabidopsis mur1*. Science 272: 1808-1810.

Histograms for complementation analysis
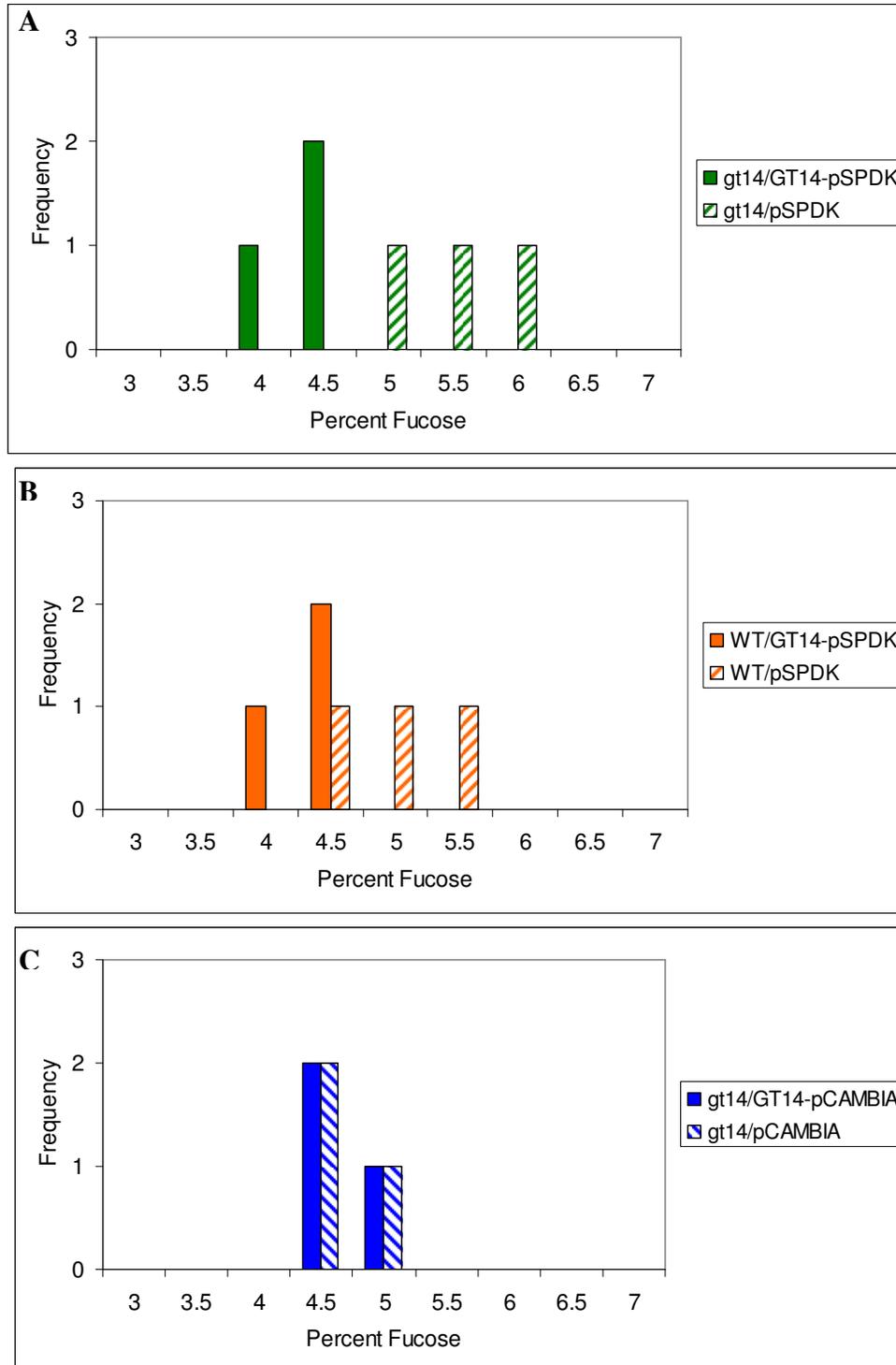
Figure 13. Histograms comparing fucose composition of T1 *gt14* and WT plants transformed with *GT14* in pCAMBIA and pSPDK vectors. None of the transformants showed significant changes relative to the control groups.
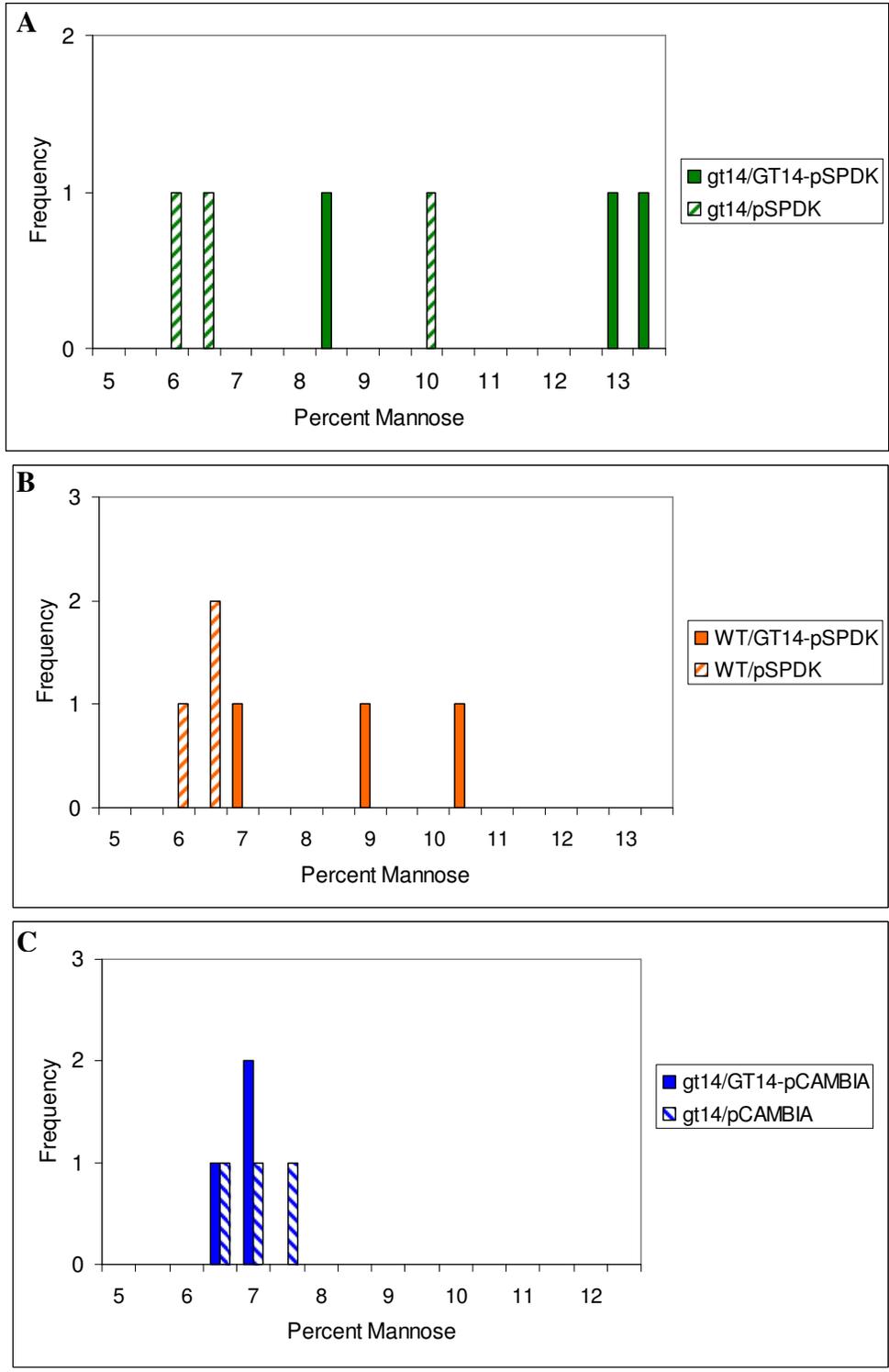
Figure 14. Histograms comparing mannose composition of T1 *gt14* and WT plants transformed with *GT14* in pCAMBIA and pSPDK vectors. Some plants in the pSPDK complemented line (**A**) may have less mannose than the control group.
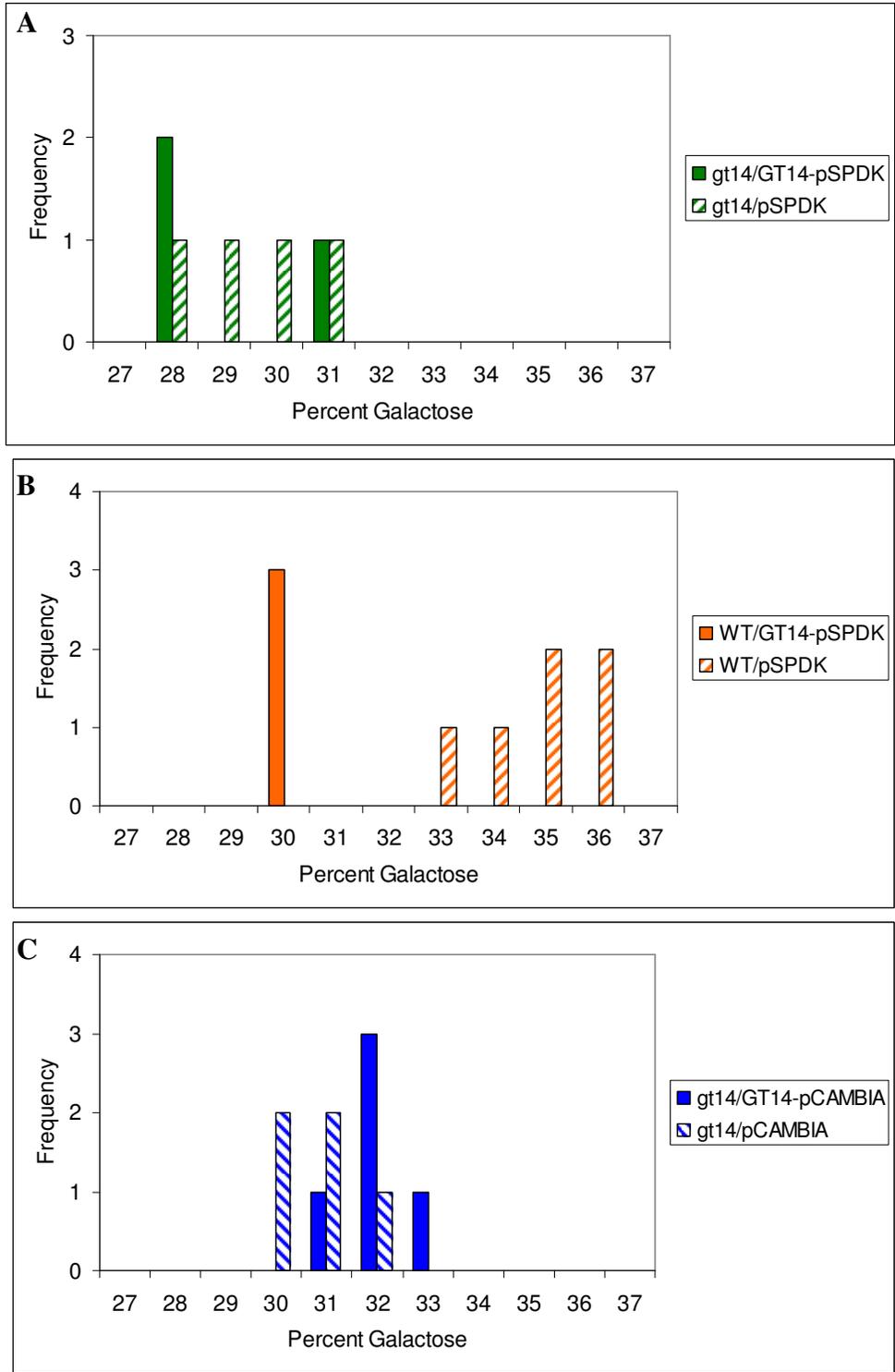
Figure 15. Histograms comparing galactose composition of T1 *gt14* and WT plants transformed with *GT14* in pCAMBIA and pSPDK vectors. No changes were observed in the complemented (**A, C**) transformants, but the overexpressed (**B**) transformants had relatively less galactose than the control plants.
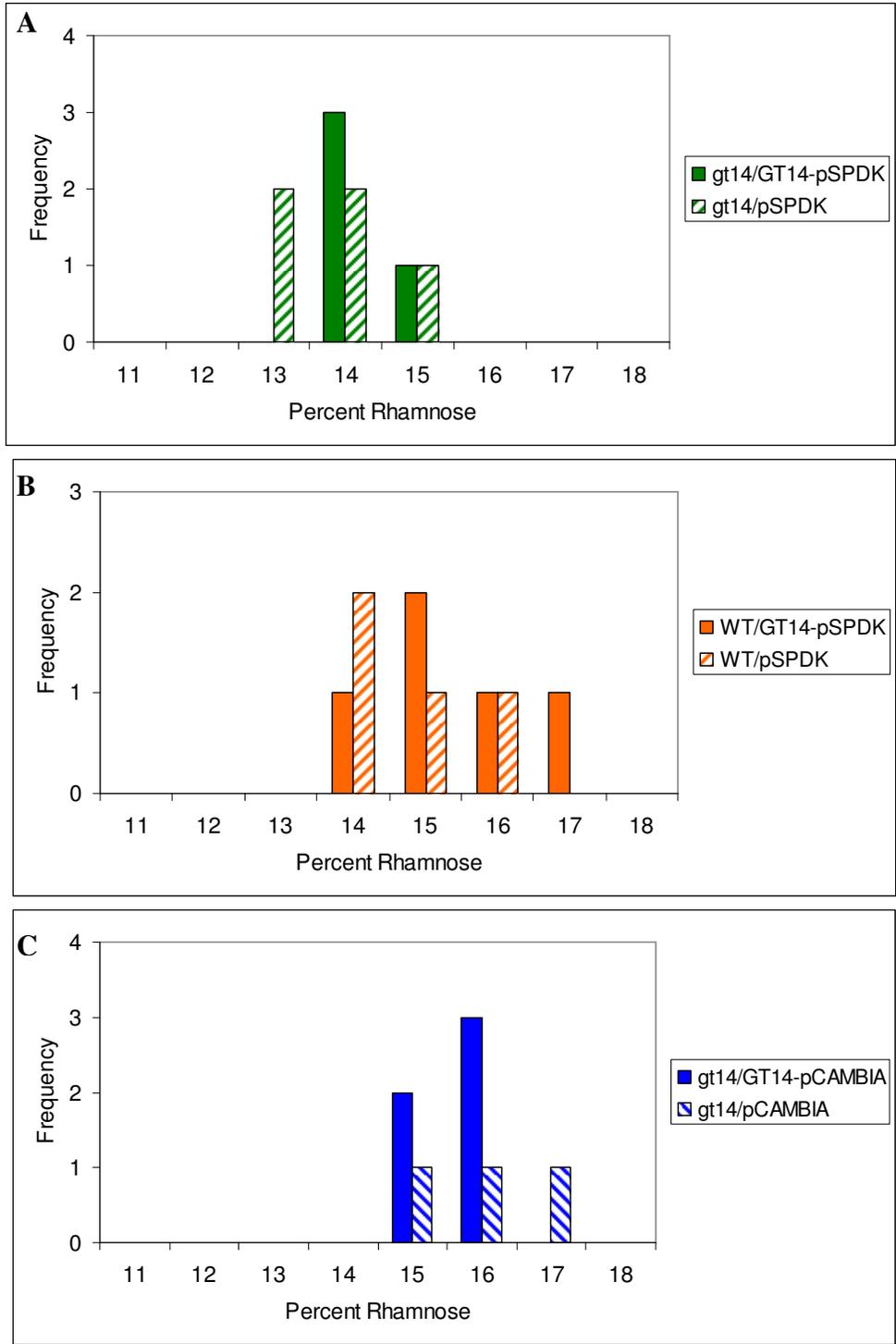
Figure 16. Histograms comparing rhamnose composition of T1 *gt14* and WT plants transformed with *GT14* in pCAMBIA and pSPDK vectors. None of the transformants showed significant changes relative to the control groups.
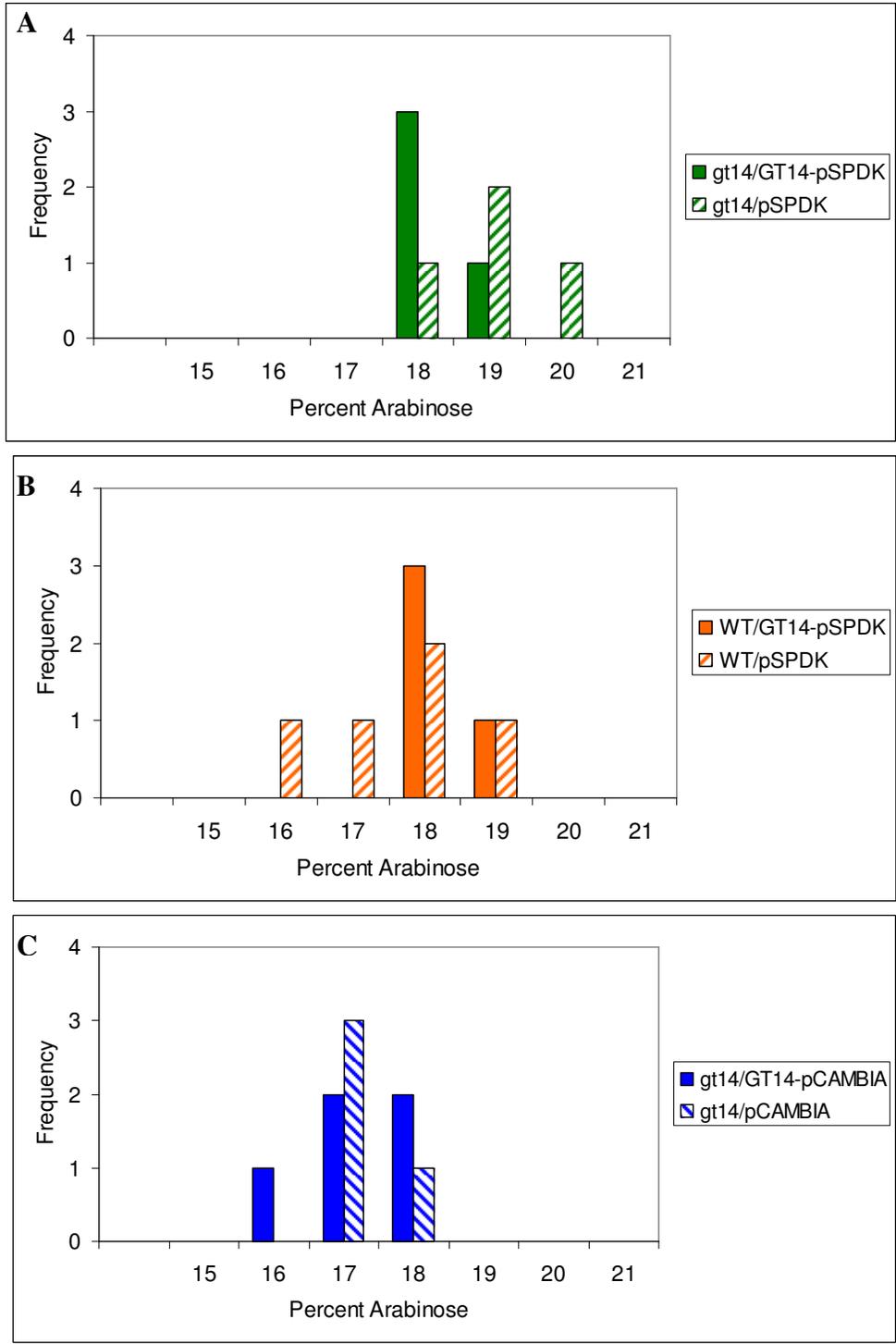
Figure 17. Histograms comparing arabinose composition of T1 *gt14* and WT plants transformed with *GT14* in pCAMBIA and pSPDK vectors. None of the transformants showed significant changes relative to the control groups.
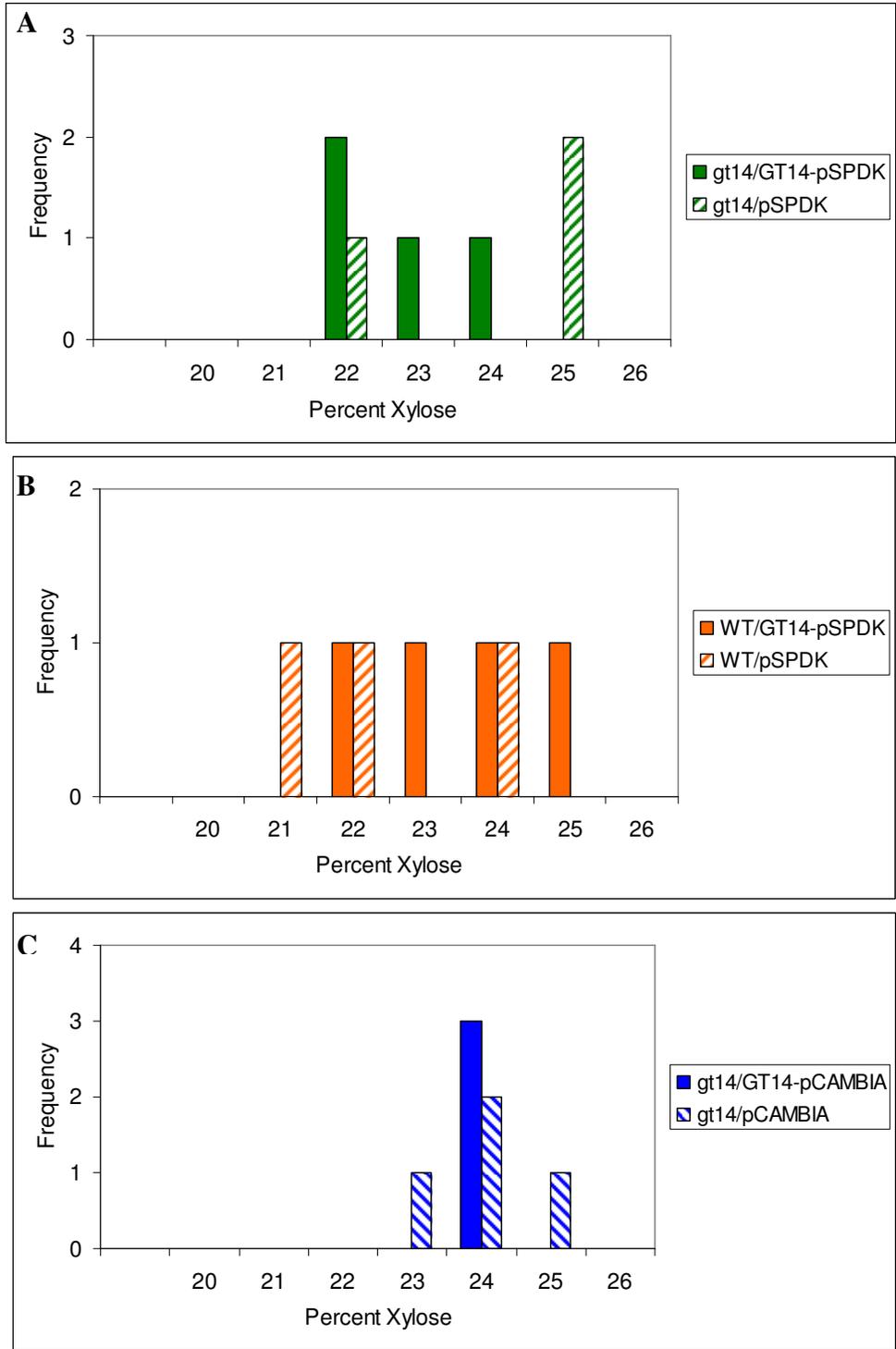
Figure 18. Histograms comparing xylose composition of T1 *gt14* and WT plants transformed with *GT14* in pCAMBIA and pSPDK vectors. None of the transformants showed significant changes relative to the control groups.

Sequence of the pSPDK1677 destination vector

5'-tccaccatgttatcacatcaatccacttgctttgaagacgtggttggaacgtcttcttttccacgatgctcctcgtgggtggggtccatctttgggaccac
tgtcggcagaggcatcttgaacgatagcctttcctttatcgcaatgatggcatttgtaggtgccaccttcctttctactgtccttttgatgaagtgacagatagctg
ggcaatggaatccgaggaggtttcccgatattaccctttgttgaaaagtctcaatagcccttggtcttctgagactgtatctttgatattcttggagtagacgaga
gtgtcgtgctccaccatgttggcaagctgctctagcattcgccattcaggctgcgcaactgttgggaagggcgatcggtgcgggcctcttcgctattacgcca
gctggcgaaaggggggatgtgctgcaaggcgattaagttgggtaacgccagggttttcccagtcacgacgttgtaaaacgacggccagtgccAAGCT
TGCATGCCTGCAGGTCAACATGGTGGAGCACGACACTCTCGTCTACTCCAAGAATATCAAAGA
TACAGTCTCAGAAGACCAGAGGGCTATTGAGACTTTTCAACAAAGGGTAATATCGGGAAACC
TCCTCGGATTCCATTGCCCAGCTATCTGTCACTTCATCGAAAGGACAGTAGAAAAGGAAGATG
GCTTCTACAAATGCCATCATTGCGATAAAGGAAAGGCTATCGTTCAAGATGCCTCTACCGACA
GTGGTCCCAAAGATGGACCCCCACCCACGAGGAACATCGTGGAAAAAGAAGACGTTCCAACC
ACGTCTTCAAAGCAAGTGGATTGATGTGATGGTCAACATGGTGGAGCACGACACTCTCGTCTA
CTCCAAGAATATCAAAGATACAGTCTCAGAAGACCAGAGGGCTATTGAGACTTTTCAACAAA
GGGTAATATCGGGAAACCTCCTCGGATTCCATTGCCCAGCTATCTGTCACTTCATCGAAAGGA
CAGTAGAAAAGGAAGATGGCTTCTACAAATGCCATCATTGCGATAAAGGAAAGGCTATCGTT
CAAGATGCCTCTACCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAACATCGTGGA
AAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGATTGATGTGATATCTCCACTGACGT
AAGGGATGACGCACAATCCCACTATCCTTCGCAAGACCCTTCCTCTATATAAGGAAGTTCATT
TCATTTGGAGaggccttctagagcggccgcgggatccGGTACCacaagtttgtacaaaaaagctgaacgagaaacgtaaatgatataaatat
caatatattaaattagattttgcataaaaaacagactacataataactgtaaaacacaacatatccagtcactatggcggccgcattaggcaccccaggctttaca
ctttatgcttccggctcgtataatgtgtggattttgagttaggatccgtcgagattttcaggagctaaggaagctaaaATGGAGAAAAAAATCA
CTGGATATACCACCGTTGATATATCCCAATGGCATCGTAAAGAACATTTTGAGGCATTTCAGT
CAGTTGCTCAATGTACCTATAACCAGACCGTTCAGCTGGATATTACGGCCTTTTTAAAGACCG
TAAAGAAAAATAAGCACAAGTTTTATCCGGCCTTTATTCACATTCTTGCCCGCCTGATGAATG
CTCATCCGGAATTCCGTATGGCAATGAAAGACGGTGAGCTGGTGATATGGGATAGTGTTCACC
CTTGTTACACCGTTTTCCATGAGCAAACTGAAACGTTTTCATCGCTCTGGAGTGAATACCACG
ACGATTTCCGGCAGTTTCTACACATATATTCGCAAGATGTGGCGTGTTACGGTGAAAACCTGG
CCTATTTCCCTAAAGGGTTTATTGAGAATATGTTTTTCGTCTCAGCCAATCCCTGGGTGAGTTT
CACCAGTTTTGATTTAAACGTGGCCAATATGGACAACTTCTTCGCCCCCGTTTTCACCATGGGC
AAATATTATACGCAAGGCGACAAGGTGCTGATGCCGCTGGCGATTCAGGTTCATCATGCCGTT
TGTGATGGCTTCCATGTCGGCAGAATGCTTAATGAATTACAACAGTACTGCGATGAGTGGCAG
GCGGGGCGTAATCTAGAGGATCCGGCTTACTAAaagccagataacagtatgcgtatttgcgcgctgattttttgcggtataaga
atatatactgatatgtatacccgaagtatgtcaaaaagaggtatgctatgaagcagcgtattacagtgacagttgacagcgacagctatcagttgctcaaggca
tatatgatgtcaatatctccggtctggtaagcacaaccatgcagaatgaagcccgtcgtctgcgtgccgaacgctggaaagcggaaaatcaggaagggatg
gctgaggtcgcccggtttattgaaatgaacggctctttttgctgacgagaacagggggctggtgaaATGCAGTTTAAGGTTTACACCTAT
AAAAGAGAGAGCCGTTATCGTCTGTTTGTGGATGTACAGAGTGATATTATTGACACGCCCGGG
CGACGGATGGTGATCCCCCTGGCCAGTGCACGTCTGCTGTCAGATAAAGTCCCCCGTGAACTT
TACCCGGTGGTGCATATCGGGGATGAAAGCTGGCGCATGATGACCACCGATATGGCCAGTGT
GCCGGTCTCCGTTATCGGGGAAGAAGTGGCTGATCTCAGCCACCGCGAAAATGACATCAAAA
ACGCCATTAACCTGATGTTCTGGGGAATATAAatgtcaggctcccttatacacagccagtctgcaggtcgaccatagtgactg
gatatgttgtgttttacagtattatgtagtctgtttttttatgcaaaatctaatttaatatattgatatttatatcattttacgtttctcgttcagctttcttgtacaaagtggtt
CCTAGGAGATCCATCGCCACCatggtgagcaagggcgaggagctgttcaccggggtggtgcccatcctggtcgagctggacggcga
cgtaaacggccacaagttcagcgtgtccggcgagggcgagggcgatgccacctacggcaagctgaccctgaagttcatctgcaccaccggcaagctgc
ccgtgccctggcccaccctcgtgaccacccttcggctacggcctgatgtgcttcgcccgctaccccgaccacatgaagcagcacgacttcttcaagtccgcc
atgcccgaaggctacgtccaggagcgcaccatcttcttcaaggacgacggcaactacaagacccgcgccgaggtgaagttcgagggcgacaccctggt
gaaccgcatcgagctgaagggcatcgacttcaaggaggacggcaacatcctggggcacaagctggagtacaactacaacagccacaacgtctatatcat
ggccgacaagcagaagaacggcatcaaggtgaacttcaagatccgccacaacatcgaggacggcagcgtgcagctcgccgaccactaccagcagaac
accccatcggcgacggccccgtgctgctgcccgacaaccactacctgagctaccagtccgccctgagcaaagaccccaacgagaagcgcgatcacat
ggtcctgctggagttcgtgaccgccgccgggatcactctcggcatggacgagctgtacaagtgaGAGCTCggaa***gatcgttcaaacatttggca
ataaagtttcttaagattgaatcctgttgccggtcttgcgatgattatcatataatttctgttgaattacgttaagcatgtaataattaacatgtaatgcatgacgttatt
tatgagatgggtttttatgattagagtcccgcaattatacatttaatacgcgatagaaaacaaaatatagcgcgcaaactaggataaattatcgcgcgcggtgtc
atctatgttactagatc****GAATTCgtaatcatggtcatagctgtttcctgtgtgaaattgttatccgctcacaattccacacaacatacgagccggaag
cataaagtgtaaagcctggggtgcctaatgagtgagctaactcacattaattgcgttgcgctcactgcccgctttccagtcgggaaacctgtcgtgccagctg
cattaatgaatcggccaacgcgcggggagaggcggtttgcgtattggctagagcagcttgagcttggatcagattgtcgtttcccgccttcagtttaaactatc

48

agtgtttgacaggatatattggcgggtaaacctaagagaaaagagcgtttattagaataacggatatttaaaagggcgtgaaaaggtttatccgttcgtccattt
gtatgtgcatgccaaccacagggttcccctcgggatcaagaattcagtactttgatccaacccctccgctgctatagtgcagtcggcttctgacgttcagtgca
gccgtcttctgaaaacgacatgtcgcacaagtcctaagttacgcgacaggctgccgccctgccctttcctggcgtttttcttgtcgcgtgttttagtcgcataaa
gtagaatacttgcgactagaaccggagacattacgccatgaacaagagcgccgccgctggcctgctgggctatgcccgcgtcagcaccgacgaccagg
acttgaccaaccaacgggccgaactgcacgcggccggctgcaccaagctgtttttccgagaagatcaccggcaccaggcgcgaccgcccggagctggc
caggatgcttgaccacctacgccctggcgacgttgtgacagtgaccaggctagaccgcctggcccgcagcacccgcgacctactggacattgccgagcg
catccaggaggccggcgcgggcctgcgtagcctggcagagccgtgggccgacaccaccacgccggccggccgcatggtgttgaccgtgttcgccggc
attgccgagttcgagcgttccctaatcatcgaccgcacccggagcgggcgcgaggccgccaaggcccgaggcgtgaagtttggcccccgcccctaccctc
accccggcacagatcgcgcacgcccgcgagctgatcgaccaggaaggccgcaccgtgaaagaggcggctgcactgcttggcgtgcatcgctcgaccc
tgtaccgcgcacttgagcgcagcgaggaagtgacgcccaccgaggccaggcggcgcggtgccttccgtgaggacgcattgaccgaggccgacgccct
ggcggccgccgagaatgaacgccaagaggaacaagcatgaaaccgcaccaggacggccaggacgaaccgttttcattaccgaagagatcgaggcgg
agatgatcgcggccgggtacgtgttcgagccgcccgcgcacgtctcaaccgtgcggctgcatgaaatcctggccggtttgtctgatgccaagctggcggc
ctggccggccagcttggccgctgaagaaaccgagcgccgccgtctaaaaaggtgatgtgtatttgagtaaaacagcttgcgtcatgcggtcgctgcgtatat
gatgcgatgagtaaataaacaaatacgcaaggggaacgcatgaaggttatcgctgtacttaaccagaaaggcgggtcaggcaagacgaccatcgcaacc
catctagcccgcgccctgcaactcgccggggccgatgttctgttagtcgattccgatccccagggcagtgcccgcgattgggcggccgtgcgggaagatc
aaccgctaaccgttgtcggcatcgaccgcccgacgattgaccgcgacgtgaaggccatcggccggcgcgacttcgtagtgatcgacggagcgccccag
gcggcggacttggctgtgtccgcgatcaaggcagccgacttcgtgctgattccggtgcagccaagcccttacgacatatgggccaccgccgacctggtgg
agctggttaagcagcgcattgaggtcacggatggaaggctacaagcggcctttgtcgtgtcgcgggcgatcaaaggcacgcgcatcggcggtgaggttg
ccgaggcgctggccgggtacgagctgcccattcttgagtcccgtatcacgcagcgcgtgagctacccaggcactgccgccgccggcacaaccgttcttg
aatcagaacccgagggcgacgctgcccgcgaggtccaggcgctggccgctgaaattaaatcaaaactcatttgagttaatgaggtaaagagaaaatgagc
aaaagcacaaacacgctaagtgccggccgtccgagcgcacgcagcagcaaggctgcaacgttggccagcctggcagacacgccagccatgaagcgg
gtcaactttcagttgccggcggaggatcacaccaagctgaagatgtacgcggtacgccaaggcaagaccattaccgagctgctatctgaatacatcgcgca
gctaccagagtaaatgagcaaatgaataaatgagtagatgaattttagcggctaaaggaggcggcatggaaaatcaagaacaaccaggcaccgacgccg
tggaatgccccatgtgtgtgggaacgggcggttggccaggcgtaagcggctggttgtctgccggccctgcaatggcactggaaccccccaagcccgag
gaatcggcgtgacggtcgcaaaccatccggcccggtacaaatcggcgcggcgctgggtgatgacctggtggagaagttgaaggccgcgcaggccgcc
cagcggcaacgcatcgaggcagaagcacgccccggtgaatcgtggcaagcggccgctgatcgaatccgcaaagaatcccggcaaccgccggcagcc
ggtgcgccgtcgattaggaagccgcccaagggcgacgagcaaccagattttttcgttccgatgctctatgacgtgggcacccgcgatagtcgcagcatcat
ggacgtggccgttttccgtctgtcgaagcgtgaccgacgagctggcgaggtgatccgctacgagcttccagacgggcacgtagaggtttccgcagggcc
ggccggcatggccagtgtgtgggattacgacctggtactgatggcggtttcccatctaaccgaatccatgaaccgataccgggaagggaagggagacaa
gcccggccgcgtgttccgtccacacgttgcggacgtactcaagttctgccggcgagccgatggcggaaagcagaaagacgacctggtagaaacctgcat
tcggttaaacaccacgcacgttgccatgcagcgtacgaagaaggccaagaacggccgcctggtgacggtatccgagggtgaagccttgattagccgcta
caagatcgtaaagagcgaaaccgggcggccggagtacatcgagatcgagctagctgattggatgtaccgcgagatcacagaaggcaagaacccggac
gtgctgacggttcaccccgattacttttttgatcgatcccggcatcggccgtttctctaccgcctggcacgccgcgccgcaggcaaggcagaagccagatg
gttgttcaagacgatctacgaacgcagtggcagcgccggagagttcaagaagttctgtttcaccgtgcgcaagctgatcgggtcaaatgacctgccggagt
acgatttgaaggaggaggcggggcaggctggcccgatcctagtcatgcgctaccgcaacctgatcgagggcgaagcatccgccggttcctaatgtacgg
agcagatgctagggcaaattgccctagcaggggaaaaaggtcgaaaaggtctctttcctgtggatagcacgtacattgggaacccaaagccgtacattggg
aaccggaacccgtacattgggaacccaaagccgtacattgggaaccggtcacacatgtaagtgactgatataaaagagaaaaaaggcgattttttccgccta
aaactctttaaaacttattaaaactcttaaaacccgcctggcctgtgcataactgtctggccagcgcacagccgaagagctgcaaaaagcgcctacccttcgg
tcgctgcgctccctacgccccgccgcttcgcgtcggcctatcgcggccgctggccgctcaaaaatggctggcctacggccaggcaatctaccagggcgc
ggacaagccgcgccgtcgccactcgaccgccggcgcccacatcaaggcaccctgcctcgcgcgtttcggtgatgacggtgaaaacctctgacacatgca
gctcccggagacggtcacagcttgtctgtaagcggatgccgggagcagacaagcccgtcaggcgcgtcagcgggtgttggcgggtgtcggggcgca
gccatgacccagtcacgtagcgatagcggagtgtatactggcttaactatgcggcatcagagcagattgtactgagagtgcaccatatgcggtgtgaaatac
cgcacagatgcgtaaggagaaaataccgcatcaggcgctcttccgcttcctcgctcactgactcgctgcgctcggtcgttcggctgcggcgagcggtatca
gctcactcaaaggcggtaatacggttatccacagaatcaggggataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggccaggaaccgtaa
aaaggccgcgttgctggcgtttttccataggctccgcccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgacaggacta
taaagataccaggcgtttccccctggaagctccctcgtgcgctctcctgttccgaccctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgt
ggcgctttctcatagctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggctgtgtgcacgaaccccccgttcagcccgaccgctgc
gccttatccggtaactatcgtcttgagtccaacccggtaagacacgacttatcgccactggcagcagccactggtaacaggattagcagagcgaggtatgta
ggcggtgctacagagttcttgaagtggtggcctaactacggctacactagaaggacagtatttggtatctgcgctctgctgaagccagttaccttcggaaaaa
gagttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggtttttttgtttgcaagcagcagattacgcgcagaaaaaaaggatctcaagaag
atcctttgatcttttctacggggtctgacgctcagtggaacgaaaactcacgttaagggattttggtcatgcatgatatatctcccaatttgtgtagggcttattatg
cacgcttaaaaataataaaagcagacttgacctgatagtttggctgtgagcaattatgtgcttagtgcatctaatcgcttgagttaacgccggcgaagcggcgt
cggcttgaacgaatttctagctagacattatttgccgactaccttggtgatctcgcctttcacgtagtggacaaattcttccaactgatctgcgcgcgaggccaa
gcgatcttcttcttgtccaagataagcctgtctagcttcaagtatgacgggctgatactgggccggcaggcgctccattgcccagtcggcagcgacatcttc
ggcgcgattttgccggttactgcgctgtaccaaatgcgggacaacgtaagcactacatttcgctcatcgccagcccagtcgggcggcgagttccatagcgtt
aaggtttcatttagcgcctcaaatagatcctgttcaggaaccggatcaaagagttcctccgccgctggacctaccaaggcaacgctatgttctcttgcttttgtc

agcaagatagccagatcaatgtcgatcgtggctggctcgaagatacctgcaagaatgtcattgcgctgccattctccaaattgcagttcgcgcttagctggat
aacgccacggaatgatgtcgtcgtgcacaacaatggtgacttctacagcgcggagaatctcgctctctccaggggaagccgaagtttccaaaaggtcgttg
atcaaagctcgccgcgttgtttcatcaagccttacggtcaccgtaaccagcaaatcaatatcactgtgtggcttcaggccgccatccactgcggagccgtaca
aatgtacggccagcaacgtcggttcgagatggcgctcgatgacgccaactacctctgatagttgagtcgatacttcggcgatcaccgcttcccccatgatgtt
taactttgtttagggcgactgccctgctgcgtaacatcgttgctgctccataacatcaaacatcgacccacggcgtaacgcgcttgctgcttggatgcccgag
gcatagactgtaccccaaaaaaacatgtcataacaagaagccatgaaaaccgccactgcgccgttaccaccgctgcgttcggtcaaggttctggaccagtt
gcgtgacggcagttacgctacttgcattacagcttacgaaccgaacgaggcttatgtccactgggttcgtgcccgaattgatcacaggcagcaacgctctgt
catcgttacaatcaacatgctaccctccgcgagatcatccgtgtttcaaacccggcagcttagttgccgttcttccgaatagcatcggtaacatgagcaaagtc
tgccgccttacaacggctctcccgctgacgccgtcccggactgatgggctgcctgtatcgagtggtgattttgtgccgagctgccggtcggggagctgttgg
ctggctggtggcaggatatattgtggtgtaaacaaattgacgcttagacaacttaataacacattgcggacgtttttaatgtactgaattaacgccgaattaattc
gggggatctatagggactttaggtgatctggattttagtactggattttggtttaggaattagaaattttattgatagaagtattttacaaatacaaatacatactaa
gggtttcttatatgctcaacacgtgagcgaaaccctataagaaccctaattcccttatctgggaactactcacacattattatggagaaactcgatgtcgatcga
ctctagctagaacgaattgttaggtggcggtacttgggtcgatatcaaagtgcatcacttcttcccgtatgcccaactttgtatagagagccactgcgggatcgt
caccgtaatctgcttgcacgtagatcacataagcaccaagcgcgttggcctcatgcttgaggagattgatgagcgcggtggcaatgccctgcctccggtgct
cgccggagactgcgagatcatagatatagatctcactacgcggctgctcaaacctgggcagaacgtaagccgcgagagcgccaacaaccgcttcttggtc
gaaggcagcaagcgcgatgaatgtcttactacggagcaagttcccgaggtaatcggagtccggctgatgttgggagtaggtggctacgtctccgaactcac
gaccgaaaagatcaagagcagcccgcatggatttgacttggtcagggccgagcctacatgtgcgaatgatgcccatacttgagccacctaactttgtttttag
ggcgactgccctgctgcgtaacatcgttgctgctgcgtaccatggtcgatcgacagatctgcgaatcgagagagatagatttgtagagagagactggtgattt
cagcgtgtcctctccaaatgaaatgaacttccttatatagaggaagggtcttgcgaaggatagtgggattgtgcgtcatcccttacgtcagtggagatatcaca
tcaatccacttgctttgaagacgtggttggaacgtcttcttttccacgatgctcctcgtgggtgggggtccatctttgggaccactgtcggcagaggcatcttg
aacgatagcctttcctttatcgcaatgatggcatttgtaggtgccaccttccttttctactgtcctttgatgaagtgacagatagctgggcaatggaatccgagg
aggtttcccgatattacccctttgttgaaaagtctcaatagccctttggtcttctgagactgtatctttgatattcttggagtagacgagagtgtcgtgc-3'