

# Functional identification of an *Arabidopsis* pectin biosynthetic homogalacturonan galacturonosyltransferase

Jason D. Sterling\*<sup>†</sup>, Melani A. Atmodjo\*<sup>‡</sup>, Sarah E. Inwood\*, V. S. Kumar Kolli\*<sup>§</sup>, Heather F. Quigley\*, Michael G. Hahn\*<sup>¶</sup>, and Debra Mohnen\*<sup>¶||</sup>

\*Complex Carbohydrate Research Center and Departments of <sup>‡</sup>Biochemistry and Molecular Biology and <sup>¶</sup>Plant Biology, University of Georgia, 315 Riverbend Road, Athens, GA 30602-4712

Communicated by Christopher R. Somerville, Carnegie Institution of Washington, Stanford, CA, January 20, 2006 (received for review December 12, 2005)

Galacturonosyltransferases (GalATs) are required for the synthesis of pectin, a family of complex polysaccharides present in the cell walls of all land plants. We report the identification of a pectin GalAT (GAUT1) using peptide sequences obtained from *Arabidopsis thaliana* proteins partially purified for homogalacturonan (HG)  $\alpha$ -1,4-GalAT activity. Transient expression of GAUT1 cDNA in the human embryonic kidney cell line HEK293 yielded uridine diphosphogalacturonic acid:GalAT activity. Polyclonal antibodies generated against GAUT1 immunoadsorbed HG  $\alpha$ -1,4-GalAT activity from *Arabidopsis* solubilized membrane proteins. BLAST analysis of the *Arabidopsis* genome identified a family of 25 genes with high sequence similarity to GAUT1 and homologous genes in other dicots, in rice, and in *Physcomitrella*. Sequence alignment and phylogenetic Bayesian analysis of the *Arabidopsis* GAUT1-related gene family separates them into four related clades of GAUT and GAUT-like genes that are distinct from the other *Arabidopsis* members of glycosyltransferase family 8. The identification of GAUT1 as a HG GalAT and of the GAUT1-related gene family provides the genetic and biochemical tools required to study the function of these genes in pectin synthesis.

biosynthesis | cell wall | glycosyltransferase | polygalacturonic acid | pectic polysaccharide

Pectins are structurally complex plant cell-wall polysaccharides that contain 1,4-linked  $\alpha$ -D-galactopyranosyluronic acid residues. Galacturonic acid (GalA) is the most abundant glycosyl residue in the three types of pectin present in all plant primary walls: homogalacturonan (HG), rhamnogalacturonan I (RG-I), and rhamnogalacturonan II (RG-II) (1). Pectin accounts for  $\approx$ 35% of dicot and nongraminaceous monocot primary walls and  $\approx$ 10% of the primary wall of grasses (2). Pectins are also present in the walls of gymnosperms, pteridophytes, and bryophytes as well as *Chara*, a charophycean algae that is believed to be the closest extant relative of land plants (3).

Numerous studies show that pectins contribute to the physical and biochemical properties of the wall (1) and are required for normal plant growth and development (4). A complete understanding of pectin function requires knowledge of pectin biosynthetic enzymes and their corresponding genes. Although the enzyme activities of proteins encoded by some pectin biosynthetic genes, particularly those involved in the synthesis of sugar nucleotides, have been elucidated (5–9), the activities of putative pectin biosynthetic glycosyltransferases [e.g., QUA1 (10) and NpGUT1 (11)] have not been definitively demonstrated.

HG is a polymer of  $\alpha$ -1,4-linked GalA that accounts for  $\approx$ 65% of pectin.  $\alpha$ -1,4-Galacturonosyltransferase (GalAT) activity has been identified in numerous plants (12) and shown to be membrane-bound in all species studied. Work in pea localized HG:GalAT activity to the luminal side of Golgi vesicles (13), the same location as pectin synthesis (14, 15). The most extensive study of HG:GalAT activity was done in tobacco (16, 17), and

characteristics of the tobacco enzyme are comparable with the activity described in other plant species.

*In vitro*, GalAT transfers GalA from uridine diphospho-GalA (UDP-GalA) onto endogenous acceptors to produce pectic products with molecular masses of 100 kDa to  $>$ 500 kDa (13, 16). Establishment of conditions to recover detergent-solubilized GalAT activity from membrane fractions (17) and *in vitro* studies using radiolabeled substrate (17) or fluorescently tagged (18, 19) acceptors established that, *in vitro*, GalAT preferentially transfers GalA onto the nonreducing end (20) of HG oligosaccharide acceptors [oligogalacturonides (OGAs)] of a degree of polymerization (DP)  $>$ 9 (17, 18), although OGA acceptors as small as a trimer can be used (18, 19). Polymeric pectins, such as poly-GalA and pectin, are less favorable substrates (21). Membrane-permeabilized GalAT activity from pumpkin yielded a population of OGAs elongated by up to five galacturonosyl residues (19), whereas the solubilized petunia enzyme added up to 27 galacturonosyl residues onto the OGA acceptors (18). Clarification of the mode of action of GalAT(s) and the mechanism of HG synthesis requires access to purified or recombinantly expressed enzyme(s).

A protein-purification approach was taken to identify an *Arabidopsis thaliana* gene encoding HG:GalAT, because no gene encoding an enzymatically verified GalAT has been identified in any organism. GalAT was partially purified from *Arabidopsis* suspension-cultured cells and bioinformatics together with peptide sequence data were used to identify two putative GalATs. Functional characterization of the protein encoded by one of these genes (At3g61130), by using numerous biochemical methods, provides compelling evidence that this protein is a HG:GalAT (16). We therefore named this gene galacturonosyltransferase 1 (GAUT1).

## Results and Discussion

**Identification of GAUT1.** Solubilized membrane proteins were isolated from log-phase suspension-cultured *A. thaliana* cells (22). GalAT activity was partially purified by sequential SP-Sepharose, Reactive yellow 3, and repetitive UDP-agarose chromatography (see lane 5 in Fig. 4, and Table 1, which are published as supporting information on the PNAS web site). The

Conflict of interest statement: No conflicts declared.

Abbreviations: DP, degree of polymerization; EPG, exopolysaccharonase; GalA, galacturonic acid; GalAT, galacturonosyltransferase; GATL, GAUT-like; HA, hemagglutinin; HG, homogalacturonan; HEK, human embryonic kidney; OGA, oligogalacturonide; RG-I, rhamnogalacturonan I; RG-II, rhamnogalacturonan II; UDP-GalA, uridine diphospho-GalA.

Data deposition: The sequence reported in this paper has been deposited in the GenBank database (accession no. DQ370437).

<sup>†</sup>Present address: Bioplex Therapeutics, 158 Credle Street, Pittsboro, NC 27312.

<sup>§</sup>Present address: Windber Research Institute, 620 Seventh Street, Windber, PA 15963.

<sup>||</sup>To whom correspondence should be addressed. E-mail: dmohnen@ccrc.uga.edu.

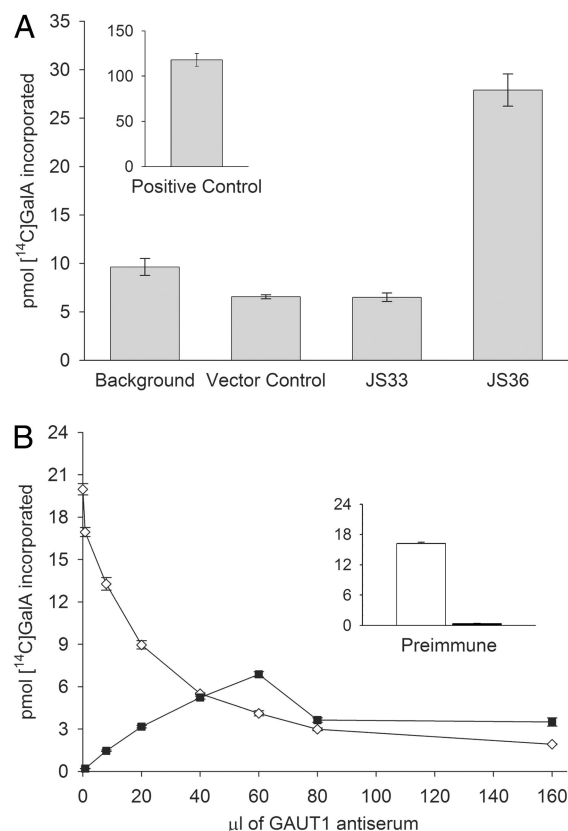
© 2006 by The National Academy of Sciences of the USA

GalAT fraction was treated with trypsin, and resulting peptides were sequenced by liquid chromatography tandem MS. The peptide sequences obtained were screened against the *Arabidopsis* genome, revealing two related proteins (JS33 and JS36) that correspond to coding sequences At2g38650 and At3g61130, respectively (see Fig. 5A and Table 2, which are published as supporting information on the PNAS web site). Both sequences are predicted to encode glycosyltransferases with type II membrane topology (23), C-terminal catalytic domains with consensus sequences clustered in CAZy glycosyltransferase family 8 (24), and to have basic isoelectric points. These characteristics are consistent with the known *in vitro* biochemical properties of GalATs (17, 20). The predicted catalytic domains also contain Dx/D motifs, a conserved motif present in many well characterized glycosyltransferase families (25) and involved in the coordination of divalent cations (26). RT-PCR showed that transcripts from both genes are present in *Arabidopsis* flowers, roots, stems, and leaves (see Fig. 6, which is published as supporting information on the PNAS web site).

cDNA constructs containing N-terminally truncated versions of JS33 (JS33 $\Delta$ 1-43) and JS36 (JS36 $\Delta$ 1-41) downstream of a *Trypanosoma cruzi* mannosidase signal sequence, a polyhistidine tag, and two copies of a hemagglutinin (HA) epitope in the mammalian expression vector pEAK10 were generated (Fig. 5B) and used to transiently transfect human embryonic kidney (HEK)293 cells. Growth media from transiently transfected cells were treated with anti-HA antibodies bound to protein A-Sepharose. The protein immunoprecipitated from the JS36 $\Delta$ 1-41-transfected cells catalyzed the incorporation of [<sup>14</sup>C]GalA from UDP-[<sup>14</sup>C]GalA onto HG OGA acceptors (Fig. 1A). No GalAT activity was recovered from immunoprecipitates of the JS33 $\Delta$ 1-43 or empty vector-transfected cells. These results provided evidence that JS36 encoded a GalAT that catalyzes the transfer of GalA from UDP-GalA onto HG, thus, we named JS36 GAUT1.

**Immunoabsorption of GalAT Activity from Partially Purified *Arabidopsis* SP-Fraction by Using GAUT1 Antiserum.** To confirm that GAUT1 was involved in HG synthesis, GAUT1 antiserum was produced and used in an immunoprecipitation assay to deplete GalAT activity present in partially purified protein preparations from *Arabidopsis* suspension-cultured cells. The GAUT1 antiserum contained antibodies that recognized native (by immunoabsorption) and denatured (by Western analysis) recombinantly expressed truncated GAUT1 (see Fig. 7, which is published as supporting information on the PNAS web site) and an  $\approx$ 60-kDa protein present in the partially purified *Arabidopsis* SP-Sepharose fraction. The GAUT1 antiserum immunodepleted GalAT activity from the *Arabidopsis* SP-Sepharose fraction (Fig. 1B, open diamonds). Furthermore, GalAT activity was immunoabsorbed by the conjugated GAUT1 antiserum in a dose-dependent manner (Fig. 1B black boxes), confirming that GAUT1 is a GalAT. Western analysis of the depleted fraction using increasing amounts of GAUT1 antiserum (see Fig. 8A, which is published as supporting information on the PNAS web site) demonstrated the progressive disappearance of the 60-kDa protein band from the SP-Sepharose fraction. Conversely, the 60-kDa band appeared in the immunoabsorbed fractions with increasing amounts of GAUT1 antiserum (Fig. 8B).

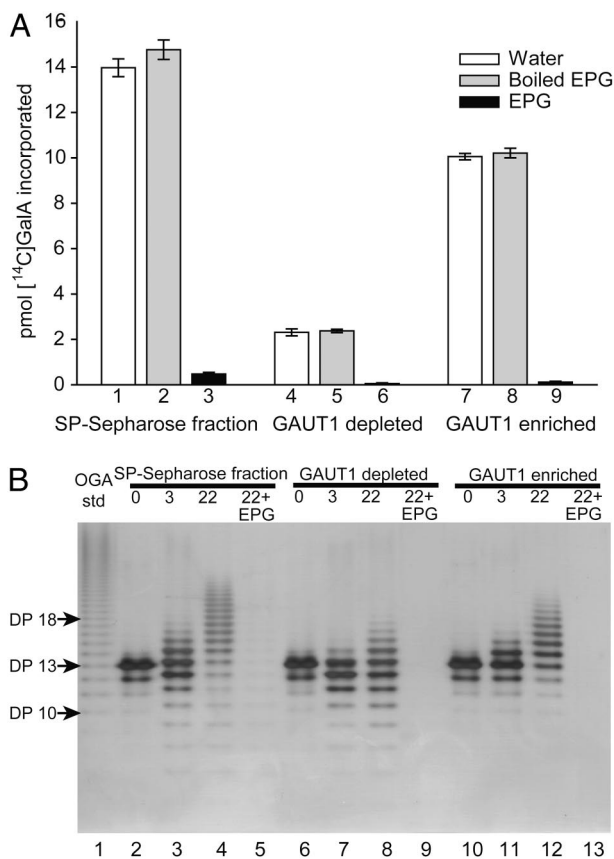
The ability of anti-GAUT1 immunoprecipitates to elongate OGAs was established by testing the sensitivity of the OGA products to cleavage by a HG-specific exopolysaccharuronase (EPG) (Fig. 2A and B). The bulk (98.9%) of the product synthesized by the anti-GAUT1 immunoprecipitates was cleaved by treatment with EPG (Fig. 2A, compare columns 7 and 8 with column 9), demonstrating that the synthesized products are HG. Furthermore, polyacrylamide-gel electrophoresis of the products generated by reacting anti-GAUT1 immunoprecipitate with



**Fig. 1.** GAUT1 has GalAT activity. (A) GalAT activity in equal amounts of media from HEK293 cells transiently transfected with the JS33 $\Delta$ 1-43 (JS33) or JS36 $\Delta$ 1-41 (JS36) cDNA constructs or empty vector control and immunoabsorbed with anti-HA antibodies conjugated to protein A-Sepharose. Solubilized *Arabidopsis* protein with GalAT activity (Inset) was used as a positive control. The average of time-0 reactions (background) is shown. Data are the average ( $\pm$ SE) of duplicate samples from 60-min reactions. Data were analyzed for significance compared with empty vector controls by using a two-sample, one-tailed Student's *t* test and an *a priori*  $\alpha$  of 0.05 [95% confidence;  $t_{JS36} = 10.284$  ( $P \approx 0.005$ );  $t_{JS33} = 0.180$  ( $P > 0.25$ )]. Comparable results were obtained in three separate experiments. (B) Immunoabsorption of GalAT activity from *Arabidopsis* solubilized membrane proteins by using GAUT1 antiserum. Solubilized *Arabidopsis* membrane proteins partially purified by SP-Sepharose chromatography were incubated with increasing amounts of GAUT1 antiserum-coated Dynabeads, and proteins immunoabsorbed by anti-GAUT1 antibodies were magnetically separated from the SP-Sepharose fraction. GalAT activity was measured in both anti-GAUT1 immunoabsorbed (black boxes) and depleted (open diamonds) fractions and compared with similar fractions obtained by using preimmune serum (Inset). Open bar, preimmune serum immunodepleted fraction; filled bar, preimmune serum immunoabsorbed fraction.

UDP-GalA and trideca-GalA revealed the appearance of OGAs of increasing size (14-mer to 21-mer) in a time-dependent fashion (compare lanes 10–12 in Fig. 2B). Conversely, only a slight increase in the size of the OGAs was observed in the immunodepleted fraction (Fig. 2B, lanes 7 and 8). The sensitivity of the elongated OGAs made by anti-GAUT1 immunoprecipitates to treatment with EPG (Fig. 2B, lane 13) clearly demonstrates that the GAUT1 antiserum immunoabsorbs HG:GalAT activity.

**GAUT1 and JS33 Are Part of a Multigene Family in *Arabidopsis*.** BLAST analysis of GAUT1 amino acid coding region against the *Arabidopsis* genome identified JS33 and 13 additional coding sequences with 36–68% amino acid sequence identity and 56–84% sequence similarity to GAUT1 (genes *GAUT 1-15*; see Table 3,



**Fig. 2.** Characterization of products made by anti-GAUT1 immunoadsorbed protein. (A) Sensitivity of synthesized product to cleavage by EPG. Products synthesized during 2-h reactions containing buffer, UDP- $^{14}\text{C}$ GalA, OGA acceptors, and the *Arabidopsis* SP-Sepharose fraction, the SP-Sepharose fraction after immunodepletion of GAUT1 (GAUT1 depleted), or the anti-GAUT1 immunoadsorbed material (GAUT1 enriched). Each fraction was incubated overnight with water, boiled EPG, or native EPG. Radiolabeled products recovered by using the filter assay are shown. Note that small oligomers (e.g., monomer and dimer) do not bind to the filters. (B) Separation of nonradioactive GalAT reaction products by electrophoresis on a 30% polyacrylamide gel. Lanes 2–4 show the products recovered after incubation of SP-Sepharose purified solubilized *Arabidopsis* membrane proteins with UDP-GalA and OGAs enriched for a DP 13 for 0, 3, and 22 h at 30°C. Lanes 6–8 show the same series of incubation times with the SP-Sepharose fraction after immunodepletion with GAUT1 antiserum. Lanes 10–12 show similar series with the product synthesized by GAUT1-immunoadsorbed protein from the SP-Sepharose fraction. Lanes 5, 9, and 13 show product recovered after digestion of the respective 22-h reaction products with EPG. Lane 1 shows 0.1  $\mu\text{g}$  of OGA standard of DP 7–23.

which is published as supporting information on the PNAS web site). An additional 10 sequences with reduced sequence identity to GAUT1 (23–29% identity and 42–53% similarity) were also identified and named the GalAT-like (*GATL*) genes (*GATL 1–10*). We suggest that all 25 *Arabidopsis* genes represent putative GalATs and have grouped them into a GAUT1-related gene superfamily (Table 3).

Sequence alignment of the 25 proteins by using CLUSTALX (27) revealed a high degree of alignment (see the first 25 entries in Fig. 9, which is published as supporting information on the PNAS web site) and the presence of several highly conserved domains (Fig. 3A) in addition to the DxD motif (25). The conserved GAUT1-related superfamily motif [H-[FWY]-[DNS]-G-x(2)-K-P-W-x(2)-[ILM]-[ADGS] is unique to the entire GAUT1-related superfamily. The *Arabidopsis* GAUT1-related superfamily members and orthologous proteins from rice, tobacco, and

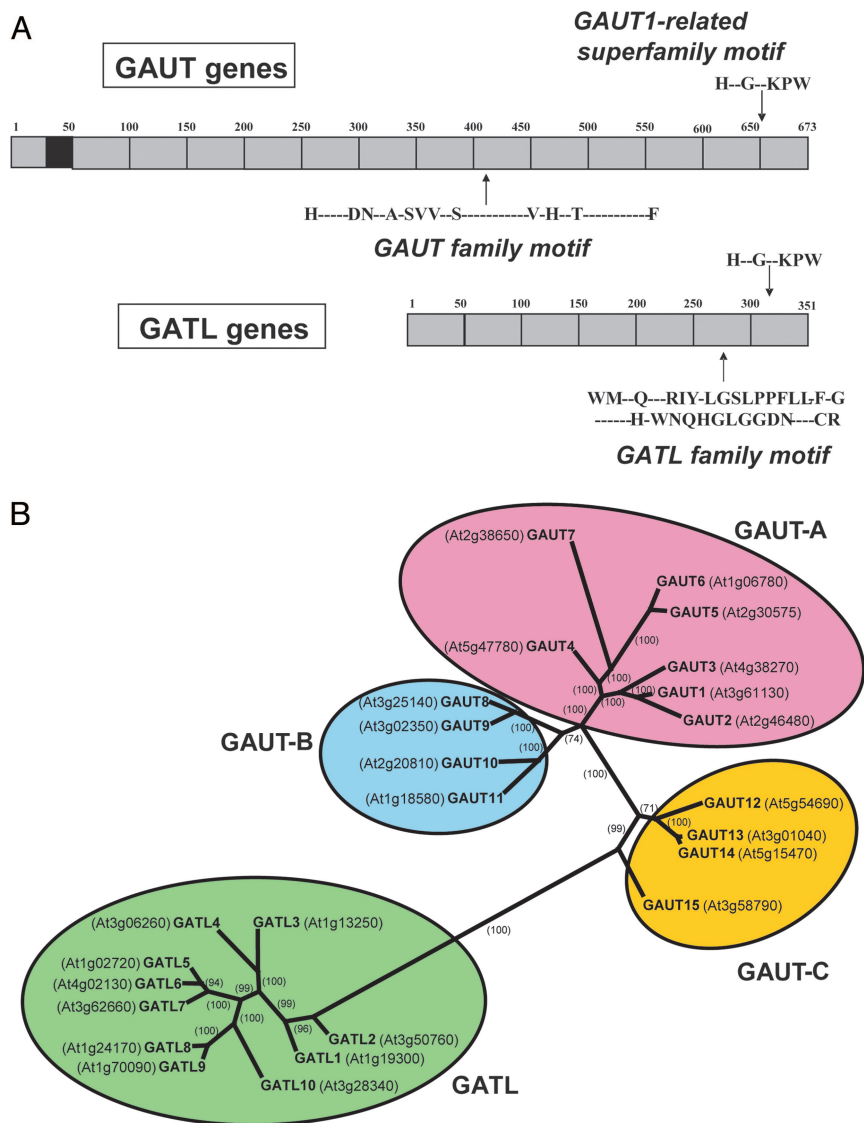
chickpea were the only sequences identified when this motif was used in a BLAST search of the SWISSPROT, TrEMBL, and BDP databases. Amino acid motifs were also identified that specifically discriminate the GAUT and GATL subclasses within the GAUT1-related superfamily. The GAUT family motif (H-x(2)-[ILV]-x-[ST]-D-N-[IV]-[IL]-A-[ASTV]-S-V-V-[AIV]-x-S-x-[AIV]-x(2)-[AS]-x(2)-[PS]-x(3)-V-[FL]-H-[ILV]-[ILV]-T-[DN]-x(2)-[NST]-x(2)-[AGP]-[IM]-x(3)-F) identifies the 15 *Arabidopsis* GAUT genes and their orthologues in other plants. The GATL motif W-M-x-[ILM]-Q-x(3,4)-R-I-Y-[DEH]-L-G-S-L-P-P-F-L-L-[IV]-F-[AGS]-G-x-[IV]-x-[AP]-[IV]-[DENS]-H-[QR]-W-N-Q-H-G-L-G-G-D-N-[FILV]-x-[GS]-x-C-R identifies the 10 *Arabidopsis* GATL genes and their orthologues.

GAUT genes are all predicted to encode proteins with molecular masses between 61 and 78 kDa (Table 3). Most of the GAUT genes are likely to encode type II membrane proteins (23) that contain a putative transmembrane domain in their hyper-variable N-terminal region (Fig. 3A). In agreement with this predicted topology, GAUT1 antiserum immunoadsorbs HG:GalAT activity, and such activity has been shown in pea to localize to the Golgi lumen (13). Furthermore, *Arabidopsis* GAUT1 protein has been localized to the Golgi (28). Three of the GAUT proteins (GAUT 3, 4, and 5) contain an N-terminal signal peptide rather than a transmembrane domain (Table 3). GAUT2 is the only member of the GAUT family that is predicted to contain no N-terminal transmembrane domain and no signal peptide. The GATL proteins have molecular masses between 39 and 44 kDa and are predicted to contain only a signal peptide at their N termini, suggesting that these proteins are not integral membrane proteins but are processed into the secretory pathway, a location consistent with a function in pectin synthesis.

The 25 proteins in the GAUT1-related superfamily belong to CAZy glycosyltransferase family 8 (24), which contains a total of 40 *Arabidopsis* proteins. We suggest that a splitting of CAZy family 8 be considered, based on functional and sequence-based analyses. Many of the current family-8 proteins are functionally annotated as being galactosyl- or glucosyltransferases, activities that are clearly distinct from the GalAT activity identified for GAUT1 and attributed to the other GAUT1-related proteins discussed here. Multiple sequence alignments of the *Arabidopsis* family-8 proteins revealed only a single region of sequence similarity across all 40 *Arabidopsis* proteins centered about the DxD motif. Even the DxD region common to the 40 *Arabidopsis* family-8 proteins contains differences that permit a distinction between the GAUT1-related family and the other family-8 proteins. For example, the GAUT1-related genes have a D-[DHS]-DxxxxxD motif, whereas the equivalent motif in the other family-8 proteins is D-[AG]-D that lacks the final D (Fig. 9). Furthermore, members of the GAUT1-related family have conserved amino acid motifs not found in any of the other *Arabidopsis* family-8 proteins (Fig. 3A). We believe that splitting family 8 as we suggest allows for evolutionarily more meaningful comparisons among this group of proteins.

**Phylogenetic Analysis of the GAUT1-Related Superfamily.** Phylogenetic analysis of the 25 GAUT1-related superfamily members show that they cluster into four distinct clades, with highly significant clade credibility values (Fig. 3B). The GAUT proteins cluster into three clades: GAUT-A (GAUT1–7), GAUT-B (GAUT8–12), and GAUT-C (GAUT12–15). All members of the GATL family cluster tightly into a distinct clade that is most closely related to GAUT15.

Our phylogenetic analysis of these proteins differs in significant respects from a published analysis (29), carried out by using CLUSTALW, particularly with respect to the phylogeny of the GAUT proteins. Our Bayesian analysis was restricted to the 25 GAUT1-related superfamily of proteins and included two proteins [GAUT11 (At1g18580) and GAUT5 (At2g30575)] not



**Fig. 3.** Characterization of the *Arabidopsis* GAUT1-related gene superfamily. (A) Schematic representation of domain structure and conserved amino acid motifs in the *Arabidopsis* GAUT and GATL proteins. Positions of conserved residues found in the GAUT family are numbered relative to GAUT1. Residues 22–44 represent a predicted transmembrane region. The positions of conserved residues in the GATL family are numbered relative to GATL1. The conserved amino acid residues within the motifs diagnostic for the GAUT1-related superfamily and the GAUT and GATL subfamilies are shown. (B) Phylogenetic analysis of the GAUT1-related superfamily in *A. thaliana*. Alignment of the complete sequences of all 25 members of the GAUT1-related superfamily was carried out with CLUSTALX (27) using parameters suggested by Hall (44). Bayesian analysis employing the program MRBAYES (45, 46) was used to infer phylogenetic relationships among the members of the superfamily and to group the protein sequences into related clades. The analysis was carried out for 100,000 generations using a mixture of amino acid transition parameter models. The phylogram presented is the majority rule tree. Percentage clade credibility values for each branch are given in parentheses.

included previously. Our analysis also specifically excluded those family-8 proteins that show no substantial regions of sequence alignment with GAUT1. Bayesian analysis yields a different set of clades within the GAUT subfamily of proteins than observed previously. QUA1 and GAUT1 fall into clearly distinct clades in our analysis, whereas they had been grouped previously into the same clade. Functional characterization of additional members of the GAUT1-related superfamily should resolve which of the two models best describe the evolutionary relationships among these proteins.

A BLAST search of the National Center for Biotechnology Information (NCBI) ([www.ncbi.nlm.nih.gov/entrez/query.fcgi](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi)) and TIGR ([www.tigr.org](http://www.tigr.org)) databases indicates the presence of several protein sequences from other plant species with significant identity to GAUT1. Phylogenetic analysis of available sequences from multiple species demonstrates that proteins with the most similarity to GAUT1 are present in soybean (TC207293), barrel medic (TC108344), tomato (TC154813), maize (TC272016), chickpea (CAB81547), and *Physcomitrella* (Contig3305). A search of the rice genome also reveals multiple proteins with high sequence similarity to members of the GAUT1-related superfamily. For example, the rice proteins Q652K2 and Q5Z7Z8 appear to be orthologues of GAUT1.

Other gene sequences from rice (e.g., BAC06990) cluster into specific clades within the *Arabidopsis* phylogenetic tree. These results strongly suggest that GAUT1 and other members of the GAUT1-related superfamily are highly conserved in vascular and nonvascular plants.

**Proposed Role in Pectin Synthesis.** Pectin consists of HG, RG-I, RG-II, and, in some tissues and species, xylogalacturonan and apiogalacturonan (1). The precise nature of the linkages between these pectic polysaccharides in the wall remains controversial (30), although the available evidence supports a linkage via the backbones of the polymers (31). Clearly, the synthesis of pectin requires the coordinated action of numerous GalATs (at least one HG:GalAT, one RG-I:GalAT, and two to three RG-II:GalATs) (12). We propose that multiple members of the GAUT1-related gene family encode GalATs involved in different aspects of pectin synthesis. Consistent with this hypothesis, a survey of the *Arabidopsis* massively parallel signature sequencing (MPSS) database, the *Arabidopsis* microarray database GENEVESTIGATOR ([www.geneinvestigator.ethz.ch](http://www.geneinvestigator.ethz.ch)) (32), and the whole-genome arrays of *Arabidopsis* (33) show that most members of the GalAT superfamily are expressed, at varying levels, in all of the major tissues of *Arabidopsis* (see Table 4,

which is published as supporting information on the PNAS web site).

Our recent analysis of walls from homozygous mutants of 12 members of the *GAUT1*-related gene family shows that mutants in 9 of the genes have significant reductions in the amount of GalA in their walls (K. Caffall and D.M., unpublished data), providing support for a function of other *GAUT1*-related genes as GalATs. Furthermore, plants carrying mutations in two members of the *Arabidopsis GAUT1*-related superfamily (*gaut8/qua1*) and (*gat1/parvus/glz1*) have been described in refs. 10, 29, 34, and 35.

*Qua1* mutant plants are dwarfed and have reduced cell adhesion and 25% reduced levels of GalA in their leaves. However, the pleiotropic effects of the mutation on glycosyltransferase activities make it difficult to specify the enzymatic function of the mutated gene (35). Our studies show that *GAUT8/QUA1* is a member of the *GAUT1*-related superfamily and support the argument that *GAUT8/QUA1* functions as a GalAT. Definitive proof will require expression of enzyme activity.

The *parvus/glz1* mutants are semisterile dwarf mutants. The neutral sugar compositions of *parvus/glz1* walls differ from wild-type walls in glycosyl residues found in RG-I. Because the levels of GalA in the *parvus/glz1* walls were not determined, it is not known whether *parvus/glz1* walls are altered in GalA content. Nevertheless, the phenotypes of plants carrying a mutation in the *GATL1/PARVUS/GLZ* and *GAUT8/QUA1* genes, together with the location of these proteins in different clades of the *GAUT1*-related superfamily, support the argument that *GAUT8/QUA1* and *GATL1/PARVUS/GLZ* function as putative GalATs involved in pectin synthesis.

The identification of *GAUT1* as a HG GalAT provides the molecular tools required to elucidate the biochemical mechanism(s) of HG and pectin synthesis. Biochemical and functional studies of *GAUT1* and other members of the *GAUT1*-related gene family are expected to increase our understanding of the biological roles of pectin in plants and may lead to enhanced agricultural productivity and development of pectin-based pharmaceuticals and industrial polymers.

## Materials and Methods

For a detailed version of this section, see *Supporting Materials and Methods*, which is published as supporting information on the PNAS web site.

**Plant and Mammalian Cell Cultures.** Cell-suspension cultures of *A. thaliana* (cv. Columbia) (22) were collected during exponential growth by filtration, washed extensively with water, and stored at  $-80^{\circ}\text{C}$  until use. HEK293 cells (Edge Biosystems, Gaithersburg, MD) were grown in bicarbonate-buffered, Dulbecco's modified Eagle's medium (Sigma), pH 7.4, containing glucose (4.5 g/liter) and supplemented with 10% (vol/vol) FBS (Sigma), 0.6 g/liter L-glutamine, 100  $\mu\text{g}/\text{ml}$  streptomycin sulfate, and 100 units/ml penicillin.

**Preparation of SP-Sepharose-Purified *Arabidopsis* Membrane Proteins.** Detergent-solubilized membrane proteins from suspension-cultured *Arabidopsis* cells (22) were loaded onto a SP-Sepharose column and bound proteins eluted by using a NaCl step gradient (see *Supporting Materials and Methods* for details). Proteins eluting with 300 and 400 mM NaCl were pooled, desalted, and stored at  $-80^{\circ}\text{C}$  until use.

**Partial Purification of *GAUT1* and Identification by Liquid Chromatography Tandem MS.** Desalted SP-Sepharose fraction was fractionated over Reactive yellow 3 and UDP-agarose columns, and the fraction most enriched for GalAT activity (see *Supporting Materials and Methods* and Table 1 for details) was treated with

sequencing grade, modified trypsin (Promega). Resulting peptides were dialyzed overnight and analyzed by using a Q-TOF2 (Waters Micromass, Milford, MA) tandem mass spectrometer with a Waters CapLC delivery system. Survey MS spectra, acquired from 450–1,700 mass units, were used to identify peptide sequences. Sequences were probed against the *Arabidopsis* genome and proteins identified by using the Mascot search engine (www.matrixscience.com) and a peptide and tandem MS tolerances of  $\pm 2$  and  $\pm 1$ , respectively.

**Cloning of JS33 (*GAUT7*) and JS36 (*GAUT1*).** A full-length cDNA encoding JS33 (GenBank accession no. AY091448) in vector pUNI51 was obtained from the *Arabidopsis* Biological Resource Center (ABRC, Columbus, Ohio). N-terminally truncated JS33 (JS33 $\Delta$ 1-43) lacking the putative transmembrane domain was generated by PCR using the sense primer FJS33Ntru (5'-ttattaccggggaattccacaatgcttcactctctggattt-3') and the antisense primer RpUNI51 (5'-ggcggctatgatctcggcggcctagaattg-3'). The primers contained restriction sites for SmaI and NotI, respectively (underscored in the primer sequences).

A version of JS36 (GenBank accession no. NM.115977) lacking its putative N-terminal transmembrane domain was generated (JS36 $\Delta$ 1-41) from total flower RNA by using RT-PCR and the gene-specific sense primer FJS36Ntru (5'-ttattaccggggaattccgaggagtgtatcatcgaattcctcaaatg-3') and antisense primer RJS36 (5'-gcccggcgcattttatcatgaagggtgcaacgacg-3'). The primers contained restriction sites for SmaI and NotI, respectively (underscored in the primer sequences).

**Expression of JS33 and JS36 in HEK293 Cells.** Purified PCR products for truncated JS33 and JS36 were cloned into pCR2.1-TOPO (Invitrogen), and SmaI/NotI fragments were subcloned into the mammalian expression vector pEAK10 (Edge Biosystems) along with an N-terminal HindIII/SmaI fragment containing a *T. cruzi* mannosidase signal sequence, a polyhistidine tag, and two copies of the HA epitope (36). HEK293 cells (37) transiently transfected with 25–30  $\mu\text{g}$  of DNA were incubated at  $37^{\circ}\text{C}$  for 48 h before harvesting.

**Production of Polyclonal Antisera.** Soluble, tetrameric multiple antigenic peptides corresponding to JS36 (*GAUT1*) amino acid sequences 448–472 (AMREYYFKADHPTSGSSNLKYRNP) coupled to a tryllysine core were synthesized at the Molecular Genetics Instrument Facility at the University of Georgia and used to produce polyclonal antibodies. Two rabbits (New Zealand White) were immunized (0.5 ml of 0.3  $\text{mg}\cdot\text{ml}^{-1}$  antigen), maintained and bled by the Polyclonal Antibody Facility at the University of Georgia.

**Magnetic Bead Immunoabsorption of GalAT Activity.** Solubilized SP-Sepharose-purified *Arabidopsis* membrane proteins containing GalAT activity were used for immunoprecipitation experiments (see *Supporting Material and Methods* for details). M-280 sheep anti-rabbit IgG-coupled Dynabeads (DynaL Biotech, Lake Success, NY) in PBS ( $6-7 \times 10^8$  beads per ml) were incubated for 2 h at  $4^{\circ}\text{C}$  with undiluted anti-*GAUT1* antiserum at a ratio of 3:1 (vol/vol). Conjugate was washed and mixed with *Arabidopsis* SP-Sepharose fraction (0–160  $\mu\text{l}$  of antiserum per 240  $\mu\text{l}$  of SP-Sepharose fraction) by rotation for 2 h at  $4^{\circ}\text{C}$ . Washed immunoprecipitates and anti-*GAUT1*-depleted supernatants were analyzed by Western analysis and by using a GalAT filter assay (38). GalAT reactions containing 30  $\mu\text{l}$  of enzyme, 50 mM Hepes, pH 7.3, 0.2 M sucrose, 0.05% (wt/vol) BSA, 25 mM KCl, 90  $\mu\text{g}$  of OGAs (DP 7–23), 6.9  $\mu\text{M}$  UDP-D-[ $^{14}\text{C}$ ]GalpA (specific activity 196  $\text{mCi}\cdot\text{mmol}^{-1}$ ) (1 Ci = 37 GBq), and 1.25 mM  $\text{MnCl}_2$  in a total reaction volume of 60  $\mu\text{l}$  were incubated for 2 h at  $30^{\circ}\text{C}$ . Reactions were terminated by the addition of 10  $\mu\text{l}$  of 0.4 M NaOH.

**Exopolysaccharide Digestion of GalAT Products.** Reaction product (total of 70  $\mu$ l) was adjusted to pH 4.5 by the addition of 10  $\mu$ l of 2 M acetic acid and 4.2  $\mu$ l of 1 M sodium acetate buffer, pH 4.2. The mixture was incubated overnight at 30°C with 2  $\mu$ l of water, a purified exopolysaccharidase (*Aspergillus tubingensis* EPG; EC 3.2.1.67, 0.54 mg/ml, 262 units/mg; 1 unit = 1  $\mu$ mol of reducing sugar produced per minute), or EPG that had been incubated at >95°C for 1 hour. The digestion reaction was terminated by addition of 23  $\mu$ l of 1 M NaOH. The final mixture was spotted onto cetylpyridinium chloride-coated filters and assayed by using the GalAT filter assay (38).

**GalAT Filter Activity Assays.** UDP-D-[<sup>14</sup>C]GalpA was synthesized as described in refs. 39 and 40. The GalAT activity assay was a modification of the procedure of ref. 16 as described in ref. 38.

**Nonradioactive GalAT PAGE Assay.** The elongation of OGA acceptors by GalAT in the presence of UDP-GalA was assayed by separating reaction products on high-percentage acrylamide gels (see below). Reaction mixtures (15  $\mu$ l) containing 0.33  $\mu$ g/ $\mu$ l OGAs (DP 13), 3 mM UDP-GalA, 1.9 mM MnCl<sub>2</sub>, 50 mM Hepes, pH 7.3, 200 mM sucrose, 25 mM KCl, 0.05% BSA, and 5  $\mu$ l of enzyme was incubated at 30°C for 3 h (unless otherwise indicated). Reactions were terminated by the addition of 700  $\mu$ l of chloroform/methanol (3:2) with vortexing and the mixture centrifuged for 5 min at 13,200  $\times$  g. The supernatant was removed and the pellet resuspended in 500  $\mu$ l of 65% ethanol by using a combination of vortexing and sonication. The mixture was centrifuged for 5 min at 13,200  $\times$  g and the supernatant removed. The pellet was air-dried for 5 min, resuspended in 50  $\mu$ l of sterile H<sub>2</sub>O, and either frozen at -20°C or analyzed directly.

Samples were separated by PAGE and visualized by alcian blue/silver nitrate staining using a modification of the procedures of refs. 41 and 42 as described by ref. 43. Samples were mixed in a 5:1 ratio with sample buffer [0.63 M Tris-Cl, pH 6.8, 0.05% phenol red, 50% (vol/vol) glycerol], loaded onto a stacking gel [5% (wt/vol) acrylamide, 0.64 M Tris, pH 6.8] and separated over a 30% acrylamide resolving gel [0.38 M Tris, pH 8.8, 30% (wt/vol) acrylamide (37.5:1 acrylamide/bis-acrylamide, wt/wt)] at 17.5 mA for 60 min. The gel was stained for 20 min with 0.2% alcian blue in 40% ethanol and washed extensively with water (three times for 20 seconds and twice for 10 min). The gel was incubated in 0.2% silver nitrate containing 0.075% formaldehyde, rinsed three times for 20 seconds with water, and developed in 4% sodium carbonate containing 0.05% formaldehyde until bands appeared. The carbonate solution was removed, and staining was terminated by addition of 5% acetic acid.

This article is dedicated to the memory of Bruce P. Wasserman, a pioneer in plant cell wall biosynthesis research. We thank Maor Bar-Peled [University of Georgia (UGA)] for the *Arabidopsis* RNA; Stefan Eberhard (UGA) for the *Arabidopsis* suspension culture; Carl Bergmann (UGA) for the exopolysaccharidase; Kelley Moremen (UGA) for the pEAK vector and the HEK293 cell line; the *Arabidopsis* Biological Resource Center and SPP Consortium: Salk/Stanford/PGECC for cDNA clone U10739; and Malcolm O'Neill, Alan Darvill, and Maor Bar-Peled for critical reading of the manuscript. This work was supported by National Research Initiative, Cooperative State Research, Education, and Extension Service, U.S. Department of Agriculture Awards 200135318-11111 and 2003-35318-15377 and, in part, by National Science Foundation Award 0090281 and Department of Energy Grant DE-FG05-93-ER20097.

- Ridley, B. L., O'Neill, M. A. & Mohnen, D. (2001) *Phytochemistry* **57**, 929–967.
- O'Neill, M., Albersheim, P. & Darvill, A. (1990) in *Methods in Plant Biochemistry*, ed. Dey, P. M. (Academic, London), Vol. 2, pp. 415–441.
- O'Neill, M. A., Ishii, T., Albersheim, P. & Darvill, A. G. (2004) *Annu. Rev. Plant Biol.* **55**, 109–139.
- Willats, W. G. T., McCartney, L., Mackie, W. & Knox, J. P. (2001) *Plant Mol. Biol.* **47**, 9–27.
- Bonin, C. P., Potter, I., Vanzin, G. F. & Reiter, W.-D. (1997) *Proc. Natl. Acad. Sci. USA* **94**, 2085–2090.
- Møhlhøj, M., Verma, R. & Reiter, W.-D. (2003) *Plant J.* **35**, 693–703.
- Burget, E. G., Verma, R., Møhlhøj, M. & Reiter, W.-D. (2003) *Plant Cell* **15**, 523–531.
- Gu, X. & Bar-Peled, M. (2004) *Plant Physiol.* **136**, 4256–4264.
- Watt, G., Leoff, C., Harper, A. D. & Bar-Peled, M. (2004) *Plant Physiol.* **134**, 1337–1346.
- Bouton, S., Leboeuf, E., Mouille, G., Leydecker, M. T., Talbotec, J., Granier, F., Lahaye, M., Höfte, H. & Truong, H.-N. (2002) *Plant Cell* **14**, 2577–2590.
- Iwai, H., Masaoka, N., Ishii, T. & Satoh, S. (2002) *Proc. Natl. Acad. Sci. USA* **99**, 16319–16324.
- Mohnen, D. (2002) in *Pectins and Their Manipulation*, eds Seymour, G. B. & Knox, J. P. (Blackwell and CRC, Oxford), pp. 52–98.
- Sterling, J., Quigley, H. F., Orellana, A. & Mohnen, D. (2001) *Plant Physiol.* **127**, 360–371.
- Northcote, D. H. & Pickett-Heaps, J. D. (1966) *Biochem. J.* **98**, 159–167.
- Moore, P. J., Swords, K. M. M., Lynch, M. A. & Staehelin, L. A. (1991) *J. Cell Biol.* **112**, 589–602.
- Doong, R. L., Liljebjelke, K., Fralish, G., Kumar, A. & Mohnen, D. (1995) *Plant Physiol.* **109**, 141–152.
- Doong, R. L. & Mohnen, D. (1998) *Plant J.* **13**, 363–374.
- Akita, K., Ishimizu, T., Tsukamoto, T., Ando, T. & Hase, S. (2002) *Plant Physiol.* **130**, 374–379.
- Ishii, T. (2002) *Plant Cell Physiol.* **43**, 1386–1389.
- Scheller, H. V., Doong, R. L., Ridley, B. L. & Mohnen, D. (1999) *Planta* **207**, 512–517.
- Takeuchi, Y. & Tsumuraya, Y. (2001) *Biosci. Biotech. Biochem.* **65**, 1519–1527.
- Guillaumie, F., Sterling, J. D., Jensen, K. J., Thomas, O. R. T. & Mohnen, D. (2003) *Carbohydr. Res.* **338**, 1951–1960.
- Reithmeier, R. A. F. & Deber, C. M. (1992) in *The Structure of Biological Membranes*, ed. Yeagle, P. (CRC, Boca Raton), pp. 337–393.
- Henrissat, B., Coutinho, P. M. & Davies, G. J. (2001) *Plant Mol. Biol.* **47**, 55–72.
- Wiggins, C. A. R. & Munro, S. (1998) *Proc. Natl. Acad. Sci. USA* **95**, 7945–7950.
- Breton, C. & Imberty, A. (1999) *Curr. Opin. Struct. Biol.* **9**, 563–571.
- Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. (1997) *Nucleic Acids Res.* **24**, 4876–4882.
- Dunkley, T. P. J., Watson, R., Griffin, J. L., Dupree, P. & Lilley, K. S. (2004) *Mol. Cell. Proteomics* **3**, 1128–1134.
- Lao, N. T., Long, D., Kiang, S., Coupland, G., Shoue, D. A., Carpita, N. C. & Kavanagh, T. A. (2003) *Plant Mol. Biol.* **53**, 687–701.
- Vincken, J.-P., Schols, H. A., Oomen, R. J. F. J., McCann, M. C., Ulvskov, P., Voragen, A. G. J. & Visser, R. G. F. (2003) *Plant Physiol.* **132**, 1781–1789.
- Nakamura, A., Furuta, H., Maeda, H., Takao, T. & Nagamatsu, Y. (2002) *Biosci. Biotechnol. Biochem.* **66**, 1301–1313.
- Zimmermann, P., Hirsch-Hoffmann, M., Hennig, L. & Gruissem, W. (2004) *Plant Physiol.* **136**, 2621–2632.
- Yamada, K., Lim, J., Dale, J. M., Chen, H., Shinn, P., Palm, C. J., Southwick, A. M., Wu, H. C., Kim, C., Nguyen, M., et al. (2003) *Science* **302**, 842–846.
- Shao, M., Zheng, H., Hu, Y., Liu, D., Jang, J.-C., Ma, H. & Huang, H. (2004) *Plant Cell Physiol.* **45**, 1453–1460.
- Orfila, C., Sørensen, S. O., Harholt, J., Geshi, N., Crombie, H., Truong, H.-N., Reid, J. S. G., Knox, J. P. & Scheller, H. V. (2005) *Planta* **22**, 613–622.
- Vandersall-Nairn, A. S., Merkle, R. K., O'Brien, K., Oelmann, T. N. & Moremen, K. W. (1998) *Glycobiology* **8**, 1183–1194.
- Jordan, M., Schallhorn, A. & Wurm, F. (1996) *Nucleic Acids Res.* **24**, 596–601.
- Sterling, J. D., Lemons, J. A., Forkner, I. F. & Mohnen, D. (2005) *Anal. Biochem.* **343**, 231–236.
- Liljebjelke, K., Adolphson, R., Baker, K., Doong, R. L. & Mohnen, D. (1995) *Anal. Biochem.* **225**, 296–304.
- Basu, S. S., Dotson, G. D. & Raetz, C. R. H. (2000) *Anal. Biochem.* **280**, 173–177.
- Corzo, J., Pérez-Galdona, R., León-Barrios, M. & Gutiérrez-Navarro, A. M. (1991) *Electrophoresis* **12**, 439–441.
- Reuhs, B. L., Carlson, R. W. & Kim, J. S. (1993) *J. Bacteriol.* **175**, 3570–3580.
- Djelino, I. (2001) *Structural Studies of Pectin*, Doctoral Dissertation (Univ. of Georgia, Athens, GA).
- Hall, B. G. (2004) *Phylogenetic Trees Made Easy: A How-To Manual* (Sinauer, Sunderland, MA), pp. 29–30.
- Ronquist, F. & Huelsenbeck, J. P. (2003) *Bioinformatics* **19**, 1572–1574.
- Huelsenbeck, J. P. & Ronquist, F. (2001) *Bioinformatics* **17**, 754–755.
- Tavares, R., Aubourg, S., Lecharny, A. & Kreis, M. (2000) *Plant Mol. Biol.* **42**, 703–717.
- Persson, S., Wei, H., Milne, J., Page, G. P. & Somerville, C. R. (2005) *Proc. Natl. Acad. Sci. USA* **102**, 8633–8638.