The analysis of global gene expression related to starch, lipid and protein composition

by

Ling Li

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Genetics

Program of Study Committee:
Eve Syrkin Wurtele, Major Professor
Martin H. Spalding
Mark E. Westgate
Dan S. Nettleton
Martha G. James
Basil J. Nikolau

Iowa State University

Ames, Iowa

2006

UMI Number: 3243820

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

TABLE OF CONTENTS

CHAPTER 1. GENERAL INTRODUCTION

INTRODUCTION

Starch, lipids and proteins are the major plant storage products. Different plant species have different relative proportions of seed stored oil, carbohydrate (often starch) and protein. Starch is the main storage reserve in grain seeds; lipids are the main storage reserve in oil seeds such as Arabidopsis; protein is highly accumulated in soybean seeds. Starch is transiently accumulated in young Arabidopsis seeds. The starch disappears later and is used as a carbon source for synthesis of fatty acid and protein (Ruuska et al., 2002). The method of conversion of starch into precursors of fatty acid biosynthesis in developing Arabidopsis seeds is not well understood (Focks and Benning, 1998). Leaves also contain storage compounds, and starch is the predominant transient reservoir of the carbon fixed during photosynthesis. Little is known about the regulatory mechanisms the plant uses to control allocation to these pathways.

The pathways that produce starch, lipid and protein are inter-related. The interconnections between fatty acid and starch metabolism are shown in Figure 1. Some key intermediates interconnect these processes, for example, hexoses and acetyl-CoA. Hexoses are precursors of starch synthesis, and product of the starch degradation. Hexoses also can be converted to pyruvate to produce acetyl-CoA and generate malonyl-CoA (Kang and Rawsthorne, 1994), for both fatty acid *de novo* synthesis and elongation of fatty acid. Here, I describe very briefly starch metabolism, in particular the *Atss3* mutant which has an increased level of starch, and an antisense-*ACLA* which has a starch excess phenotype altered in the production of cytosolic acetyl-CoA. Using mutants altered in starch accumulation, a combination of approaches such as global gene profiling, and promoter::GUS/GFP analyses and genetics, I have expanded the understanding of starch metabolism.

Starch

Starch is a storage form of carbohydrate. It is very important for human diet and industry. Arabidopsis leaf is a good model to study transient starch accumulation and degradation. Starch accumulates in leaf chloroplasts during the light time of the diurnal

cycle. In the dark, leaf starch is converted into glucose and then made into sucrose, transported to other tissues for energy production or biosynthetic conversions. Mechanisms are not yet clear for starch biosynthesis, or its degradation to soluble glucose monomers. Starch granule formation and degradation are highly regulated processes. Amylose and amylopectin are both glucose homopolymers in starch, comprise monomers joined by α-(1→4) or α-(1→6) glycosidic bonds. After the production of the glucosyl unit donor ADP-glucose, starch synthases (SS, encoded by 6 genes), branching enzymes (BE, encoded by 3 genes), debranching enzymes (DBE, encoded by 4 genes), α-amylases (AAM, encoded by 3 genes), β-amylases (BAM, encoded by 9 genes), disproportionating enzymes (DPE, encoded by 2 genes) and starch phosphorylases (PHS, encoded by 2 genes) are involved in the formation and degradation of starch. The loci for these genes are identified on the Arabidopsis starch metabolism network web site (http://www.starchmetnet.org). The significance of each gene family member is not yet understood.

Starch Biosynthesis

SS adds glucosyl units to growing chains through new α-(1→4) linkages. BE catalyzes the introduction of α-(1→6) linkages into the polymers. DBE hydrolyzes α-(1→6) linkages. DPE transfers residues between linear chains by the cleavage and reformation of α-(1→4) linkages. Synthesis of amylose may require SS only. SS, BE, DBE and DPE may be involved in the synthesis of amylopectin (Figure 1, Myers et al., 2000). There are six isoforms of starch synthase. They can be divided into two groups: granule bound starch synthase (GBSS) and soluble starch synthase (SSS). GBSS is responsible for amylose synthesis (Ball et al., 1998). The characteristics of the mutant do not always give a clear understanding of the exact function of the gene. *Atss1* mutant (with no SS1 activity) has altered-structure amylopectin, which suggests SS1 might take part in synthesizing small outer chain of amylopectin (Delvallé et al., 2005). *Atss3* mutant (loss of a SS3 function) has an increased starch biosynthesis rate, altered leaf starch structure, and an increased total SS activity (Zhang et al., 2005). It seems that a specific SS activity is increased; together, the data suggest a possible negative regulatory function of SS3. In function from mutants indicates BE1 apparently is not necessary for starch metabolism, while BE2 and BE3 are redundant and required for normal starch synthesis (Dumez et al., 2006).

Starch Degradation

Starches have a combination of phosphorolytic and hydrolytic degradation (Beck et al., 1989). PHS, DPE, DBE, AAM, and BAM may all be involved although the function of each gene is not well understood. PHS inserts phosphoryl groups from inorganic pyrophosphate into α-(1→4) bonds and releases glucose-1-phosphate. DPE releases glucose. DBE catalyzes debranching. AAM and BAM both act as α-(1→4)-specific hydrolases. AAM cleaves within a linear chain. BAM works from the non-reducing end and its degradative action stops at the branch linkage. Plastidial PHS1 may not be necessary for starch degradation in Arabidopsis leaves but may function in abiotic stress (Zeeman et al., 2004). Cytosolic PHS2 takes part in maltose metabolism (Lu et al., 2006). AAM is not necessary for transitory starch breakdown in Arabidopsis leaves (Yu et al., 2005). BAM3 and BAM8 are shown by Sparla et al. (2006) to be targeted to plastids. Chloroplastic BAM takes part in starch degradation in the dark, and BAM3 might be involved in redox-regulated starch degradation (Sparla et al., 2006).

Starch Biosynthesis and Degradation as an Intermingled Process

Some enzymes may have a role in both starch synthesis and catabolism, e.g., DBE and DPE (Myers et al., 2000). There are 4 genes encoding DBE: *ISA1*, *ISA2* and *ISA3* that encode isoamylase-type DBEs, and *PU1* that encodes a pullulanase-type DBE. *ISA1* and *PU1* are highly conserved among plants (Beatty et al., 1999). The specific roles of DBEs are unknown. Hypotheses for DBE functions are: 1) directly hydrolyze α-(1→6) linkages and contribute to the "clustered" arrangement of branches (Ball et al., 1996); 2) indirectly eliminate water-soluble polysaccharides that compete with amylopectin biosynthesis (Zeeman, 1998); 3) suppresses starch granule initiation (Burton et al., 2002). An Arabidopsis mutant lacking a chloroplastic isoamylase (ISA2) accumulates both starch and phytoglycogen and causes reduction or elimination of amylopectin (Zeeman, 1998). The glucan is soluble and highly branched. This indicates the gene may contribute to starch granule formation. In potato (Hussain et al., 2003), an ISA1/ISA2 complex is necessary for normal starch synthesis. ISA1 and ISA2 were shown to be physically associated together as a multimeric enzyme to debranch soluble glucan. In rice (Fujita et al., 2003), an antisense *ISA1* mutant had a reduction of ISA1 protein by about 94% in endosperm. The amylopectin was

changed into a water-insoluble modified amylopectin with more short chains. The PU1 activity in this mutant was similar to the WT. This suggested that ISA1 might be essential for amylopectin biosynthesis in rice endosperm. In Arabidopsis, single mutants *Atisa1* and *Atisa2* and double mutants *Atisa1*/*Atisa2* have identical phenotypes of reduced starch content and loss of the same isoamylase enzyme activity; ISA2 was suggested to be the regulatory subunit and ISA1 to be catalytic (Delatte et al., 2005). In maize (Dinges et al., 2003), a *pu1* mutant was found to have decreased rate of starch degradation, indicating that the hydrolytic activity of PU1 might contribute to starch catabolism. During kernel starch formation, PU1 functioned in glucan hydrolysis. PU1 might have a function in both starch synthesis and starch degradation. An *Atpu1* mutant loss of PU1 function has no distinguishable phenotype (Wattebled et al., 2005). But an *Atisa2*/*Atpu1* double mutant shows decreased starch content by 92%. This suggests PU1 might have an overlapped function with ISA1 or ISA2 or both. Plant with loss of ISA3 function had a starch-excess phenotype. An *Atisa3/Atpu1* double mutant loss of ISA3 and PU1 function had more severe starch excess than *Atisa3* (Delatte et al., 2006). This indicates that ISA3 and PU1 might take part in starch degradation. Critchley et al. (2001) found plastidial DPE1 is not required for starch synthesis and its primary function is involved in starch degradation in Arabidopsis. Cytosolic DPE2 is required for maltose metabolism in Arabidopsis (Lu et al., 2004).

Table S1 (Appendix C) includes a group of gene involved in or related to starch metabolism.

ATP-Citrate Lyase (ACL)

ACL generates cytosolic acetyl-CoA. ACL catalyzes following reaction:

citrate + ATP + CoA → oxaloacetate + acetyl-CoA + ADP + Pi

Acetyl-CoA is involved in multiple physiological processes that link anabolism and metabolism. Acetyl-CoA in the cytosol of plant cells is used for the synthesis of many phytochemicals including waxes, isoprenoids, and flavonoids. Two Arabidopsis cDNAs encode proteins similar to the amino and carboxy portions of human ACL. Both the Arabidopsis genes are required for ACL activity. Arabidopsis ACL is heteromeric and consists of ACLA and ACLB. The holoprotein has a heterooctomer with an A_4B_4

configuration. ACL activity and the ACLA and ACLB polypeptides are located in the cytosol. In the Arabidopsis genome, three genes encode for the ACLA subunit while two genes encode the ACLB subunit. The *ACLA* and *ACLB* mRNAs accumulation pattern is consistent with the predicted physiological needs for cytosolic acetyl-CoA, and is coordinated with the accumulation pattern of cytosolic ACCase (Fatland et al., 2002). Antisense-*ACLA* plants have reduced cuticular wax on the inflorescence stem, and contain abnormally enlarged starch granules in leaves (Fatland et al., 2005). This leads us to believe that ACL is also involved in, or at least, related to starch metabolism.

DISSERTATION ORGANIZATION

This dissertation consists of five chapters and three appendices. Chapter 1 is the general introduction; Chapter 2, 3 and 4 are three manuscripts to be submitted to Plant Journal, Plant Physiology, and Plant Cell; Chapter 5 is the general conclusion. The three manuscripts are all from my starch research project. The first is about Arabidopsis starch debranching enzymes, about the expression of the four isoforms, trying to better understand their functions and relations. The sectioned images in this part are from Dr. Hilal Ilarslan. The second is about the connection between starch metabolism and ACL utilization, trying to understand why reduced ACL activity in Arabidopsis caused more starch content in the mutant. Dr. Carol Foster and I contributed equally to this paper. I am responsible for all the figures except Figure 6-8. I am also responsible for discussion part, and the results corresponded to the figures from me. The third is about the mRNA profiling of a mutant with loss of function in gene implicated in starch synthesis. A small unique unknown gene was found with altered transcript in this mutant. Reduced transcript of this small unique gene resulted in starch excess at the end of light. To date, no paper has reported similar research on the involvement of such a small gene in starch metabolism. Appendix A describes the expression of genes related to acetyl-CoA/biotin network, including *CAC1A*, *CAC1B*, *CACT* and wax synthase or wax synthase-like gene. Appendix B is about my contribution to soybean seed composition project. Appendix C-F includes the supplemental figures and/or tables to Chapter 1-4.

LITERATURE CITED

Ball S, Guan H, James M, Myers A, Keeling P, Mouille G, Buleon A, Colonna P, Preiss J (1996) From glycogen to amylopectin: A model for the biogenesis of the plant starch granule. Cell **86:** 349–352

Ball S, Van de Wal M, Visser R (1998) Progress in understanding the biosynthesis of amylose. Trends Plant Sci **3:** 462–467

Beatty MK, Rahman A, Cao H, Woodman W, Lee M, Myers AM, James MG (1999) Purification and molecular genetic characterization of ZPU1, a pullulanase-type starch-debranching enzyme from maize. Plant Physiol **119:** 255–266

Beck E, Ziegler P (1989) Biosynthesis and degradation of starch in higher plants. Annu Rev Plant Physiol Plant Mol Biol **40:** 95–117

Burton RA, Jenner H, Carrangis L, Fahy B, Fincher GB, Hylton C, Laurie DA, Parker M, Waite D, van Wegen S, Verhoeven T, Denyer K (2002) Starch granule initiation and growth are altered in barley mutants that lack isoamylase activity. Plant J **31:** 97–112

Critchley JH, Zeeman SC, Takaha T, Smith AM, Smith SM (2001) A critical role for disproportionating enzyme in starch breakdown is revealed by a knock-out mutation in Arabidopsis. Plant J **26:** 89–100

Delatte T, Trevisan M, Parker ML, Zeeman SC (2005) Arabidopsis mutants *Atisa1* and *Atisa2* have identical phenotypes and lack the same multimeric isoamylase, which influences the branch point distribution of amylopectin during starch synthesis. Plant J **41:** 815–830

Delatte T, Umhang M, Trevisan M, Eicke S, Thorneycroft D, Smith SM, Zeeman SC (2006) Evidence for distinct mechanisms of starch granule breakdown in plants. J Biol Chem **281:** 12050–12059

Delvallé D, Dumez S, Wattebled F, Roldan I, Planchot V, Berbezy P, Colonna P, Vyas D, Chatterjee M, Ball S, Mérida Á, D'Hulst C (2005) Soluble starch synthase I: a major determinant for the synthesis of amylopectin in *Arabidopsis thaliana* leaves. Plant J **43:** 398–412

Dinges JR, Colleoni C, James MG, Myers AM (2003) Mutational analysis of the pullulanase-type debranching enzyme of maize indicates multiple functions in starch metabolism. Plant Cell **15:** 666–680

Dumez S, Wattebled F, Dauvillee D, Delvalle D, Planchot V, Ball SG, D'Hulst C (2006) Mutants of Arabidopsis lacking starch branching enzyme II substitute plastidial starch synthesis by cytoplasmic maltose accumulation. Plant Cell **18:** 2694–2709

Fatland BL, Ke J, Anderson M, Mentzen W, Cui LW, Allred C, Johnston JL, Nikolau BJ, Wurtele ES (2002) Molecular characterization of a heteromeric ATP-citrate lyase that generates cytosolic acetyl-CoA in Arabidopsis. Plant Physiol **130:** 740–756

Fatland BL, Nikolau BJ, Wurtele ES (2005) Reverse genetic characterization of cytosolic acetyl-CoA generation by ATP-citrate lyase in Arabidopsis. Plant Cell **17:** 182–203

Focks N, Benning C (1998) *Wrinkled1*: a novel low-seed-oil mutant of Arabidopsis with a deficiency in the seed-specific regulation of carbohydrate. Plant Physiol **118:** 91–101

Fujita N, Kubo A, Suh DS, Wong KS, Jane JL, Ozawa K, Takaiwa F, Inaba Y, Nakamura Y (2003) Antisense inhibition of isoamylase alters the structure of amylopectin and the physicochemical properties of starch in rice endosperm. Plant Cell Physiol **44:** 607–618

Hussain H, Mant A, Seale R, Zeeman S, Hinchliffe E, Edwards A, Hylton C, Bornemann S, Smith AM, Martin C and Bustos R (2003) Three isoforms of isoamylase contribute different catalytic properties for the debranching of potato glucans. Plant Cell **15:** 133–149

Kang F, Rawsthorne S (1994) Starch and fatty acid biosynthesis in plastids from developing embryos of oil seed rape (*Brassica napus* L.). Plant J **6:** 795–805

Lu Y, Sharkey TD (2004) The role of amylomaltase in maltose metabolism in the cytosol of photosynthetic cells. Planta **218:** 466–473

Lu Y, Steichen JM, Yao J, Sharkey TD (2006) The Role of Cytosolic -Glucan Phosphorylase in Maltose Metabolism and the Comparison of Amylomaltase in Arabidopsis and *Escherichia coli*. Plant Physiol **142:** 878–889

Myers AM, Morell MK, James MG, Ball SG (2000) Recent progress toward understanding biosynthesis of the amylopectin crystal. Plant Physiol **122:** 989–997

Ruuska SA, Girke T, Benning C, Ohlrogge JB (2002) Contrapuntal networks of gene expression during Arabidopsis seed filling. Plant Cell **14:** 1191–1206

Sparla F, Costa A, Lo Schiavo F, Pullilo P, Trost P (2006) Redox regulation of a novel plastid-targeted *β*-amylase of *Arabidopsis thaliana.* Plant Physiol **141:** 840–850

Wattebled F, Dond Y, Dumez S, Delvalle D, Planchot R, Berbezy P, Vyas D, Colonna P, Chatterjee M, Ball S, and D'Hulst C (2005) Mutants of Arabidopsis lacking a chloroplastic isoamylase accumulate phytoglycogen and an abnormal form of amylopectin. Plant Physiol **138:** 184–195

Yu T-S, Zeeman SC, Thorneycroft D, Fulton DC, Dunstan H, Lue W-L, Hegemann B, Tung S-Y, Umemoto T, Chapple A, et al (2005) *α*-Amylase is not required for breakdown of transitory starch in Arabidopsis leaves. J Biol Chem **280:** 9773–9779

Zeeman SC, Thorneycroft D, Schupp N, Chapple A, Weck M, Dunstan H, Haldimann P, Bechtold N, Smith AM, Smith SM (2004) Plastidial *α*-glucan phosphorylase is not required for starch degradation in Arabidopsis leaves but has a role in the tolerance of abiotic stress. Plant Physiol **135:** 849–858

Zeeman SC, Umemoto T, Lue WL, Au-Yeung P, Martin C, Smith AM, Chen J (1998) A mutant of Arabidopsis lacking a chloroplastic isoamylase accumulates both starch and phytoglycogen. Plant Cell **10:** 1699–1711

Zhang X, Myers AM, James MG (2005) Mutations affecting starch synthase III in Arabidopsis alter leaf starch structure and increase the rate of starch synthesis. Plant Physiol **138:** 663–674

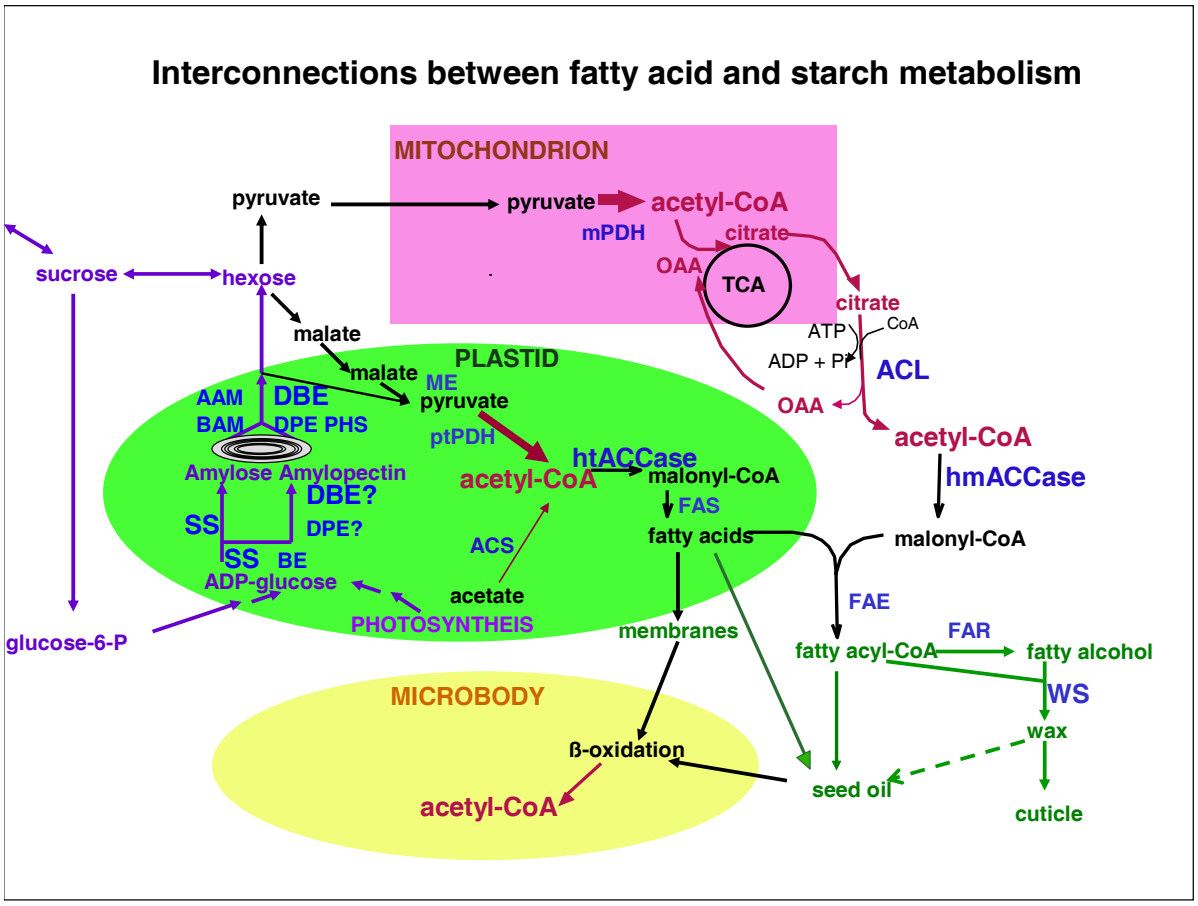


Figure 1. Interconnections between the fatty acid and starch metabolism. Hexose and acetyl-CoA interconnect the processes.

CHAPTER 2. STARCH DEBRANCHING ENZYME IN ARABIDOPSIS

A paper to be submitted to Plant Journal

Ling Li, Hilal Ilarslan, Martha G. James, Alan M. Myers, and Eve Syrkin Wurtele

ABSTRACT

Starch granule's formation and degradation are regulated in leaf chloroplasts during the light and dark photoperiods of the diurnal cycle. In addition, starch accumulates in a variety of cell types at particular stages of development and under appropriate environmental conditions. Starch debranching enzymes (DBE) are involved in both starch synthesis and degradation. This study depicted the temporal and spatial expressions of four DBEs in Arabidopsis: *ISA1*, *ISA2*, *ISA3* and *PU1*, and explores their relationship from different levels of evidence. The co-expression of *ISA1* and *ISA2* over the diurnal cycle, the high correlation of *ISA1* and *ISA2* RNA accumulation in microarray analysis, and the cellular co-expression of *ISA1* and *ISA2*, are consistent with their co-function involved in starch synthesis. *ISA2* and *ISA3* have overlapped high mRNA accumulation at the beginning of dark, which suggests their co-expression in dark involved in starch degradation. The co-expression pattern of *ISA2* and *ISA3* in tissues furthers the possibility of co-function of ISA2 and ISA3. ISA3 protein is localized in the leaf chloroplast, which supports the possibility of co-expression of *ISA2* and *ISA3*. This is the first report that gives evidence that ISA2 and ISA3 might form a protein complex. In parallel, we are exploring the starch metabolic network in the context of other metabolic processes using the MetNet platform (http://www.metnetdb.org/) to address changes in gene expression in the context of the known regulatory and metabolic network in Arabidopsis; these studies have identified genes that are candidates for regulation of starch metabolism.

INTRODUCTION

Members of a large group of structurally related enzymes referred to collectively as the α-amylase-related superfamily catalyze hydrolysis of α-glycoside bonds in glucose homopolymers. The widely distributed "α-amylase" enzymes catalyze hydrolysis of α-(1→4)

glycoside bonds in the interior of glucan chains. A class of related enzymes present in prokaryotes, the "isoamylases", catalyzes hydrolysis of α-(1→6) glycoside bonds and is known to function physiologically in catabolism of storage carbohydrates such as glycogen (Doehlert and Knutson, 1991). Another prokaryotic subfamily within the α-amylase related superfamily, the "pullulanases", are grouped with the isoamylases based on amino acid sequence similarity and catalyze hydrolysis of the same linkage, but do so with a distinct substrate specificity (Manners, 1997). Isoamylases hydrolyze α-(1→6) glycoside bonds in branched substrates such as glycogen and amylopectin but have low activity towards the glucose homopolymer pullulan, which is a linear glucan in which the linkages alternate as two α-(1→4) bonds followed by one α-(1→6) bonds. Conversely, pullulanases are active on pullulan and have only low activity on glycogen or amylopectin (Dinges et al., 2003a). Thus, the essential catalytic mechanism for α-glycoside bond hydrolysis has been conserved broadly in evolution and is used in a variety of contexts for catabolism of different glucan polymers.

In the plant kingdom there are four highly conserved members of the α-amylase-related superfamily that catalyze hydrolysis of α-(1→6) glycoside bonds (James et al., 1995; Ball et al., 1996). Three of these enzymes are most similar to the prokaryotic isoamylases and one is closest to the prokaryotic pullulanases. All four proteins are known or predicted to be present in plastids, coincident with the abundant glucan homopolymers that constitute starch granules, i.e., the predominantly α-(1→4)-linked "linear" glucan amylose, and the "branched" polymer amylopectin in which about 95% of the glycoside bonds are in the α-(1→4) configuration and about 5% in the α-(1→6) configuration. This observation suggests a role for the α-(1→4) glucosidases in breakdown of starch, particularly the amylopectin component of the granules, during periods when glucose is released from storage to supply chemical energy and precursors for metabolic interconversions. Collectively, therefore, this group of plant enzymes is referred to as starch debranching enzymes (DBE). Three of them are further classified as "isoamylase-type DBEs", and one as the "pullulanase-type DBE", to reflect their evolutionary relationships and also in some instances their known substrate specificities (Beatty et al., 1999).

The physiological functions of the DBEs are not precisely understood (Dinges et al., 2001). Pullulanase-type DBEs are well known to be involved in starch degradation during germination of monocot seeds. Likewise, in maize leaves the pullulanase-type enzyme ZmPU1

is thought to have a role in starch degradation, because loss of function mutations condition a moderate increase in the starch level at the end of the night (Beatty et al., 1999). ZmPU1 is not essential for this process, however, because extended incubation of the mutant in the dark results in complete depletion of the starch store. Similarly, in Arabidopsis, the ISA3 and PU1 isoforms have been shown by genetic analysis to be involved in leaf starch degradation during the dark phase of the diurnal cycle (Delatte et al., 2006).

Biochemical analysis of ISA1 and ISA2 mutants revealed functions distinctly different than those of ISA3 and PU1 (Hussain et al., 2003; Delatte et al., 2005; Delatte et al., 2006). In contrast to the expectation of increased starch owing to a deficiency in polymer hydrolysis, elimination of either ISA1 or ISA2 can cause a reduction in granular starch content; this reduction is most often coupled with the appearance of a soluble glycogen-like polymer that is not detected in wild type plants. This effect has been observed in endosperm of maize (James et al., 1995; Rahman et al., 1998; Dinges et al., 2001), rice (Kubo et al., 1999), and barley (Burton et al., 1999), in potato tubers (Bustos et al., 2004), and in Arabidopsis leaves (Zeeman et al., 1998b; Delatte et al., 2005; Wattebled et al., 2005). Thus, the hydrolytic activity of at least some of the isoamylase-type DBEs may contribute to the anabolic pathways leading to starch biosynthesis. Various hypotheses have been proposed to explain these observations, including both a direct role of the DBEs in producing the polymers that form starch granules and an indirect role in eliminating a metabolic competitor polymer that directs glucan away from the granules and into the soluble phase (Ball et al., 1996; Zeeman, 1998b). An alternate possibility is that the DBEs have an indirect regulatory function. Genetic analyses of pullulanase-type DBE of Maize indicate that the plant pullulanase-type DBE can also play a role in starch biosynthesis, in addition to its demonstrated function in the polymer degradative pathways (Dinges et al., 2003a; Dinges et al., 2003b).

Taken together, these observations suggest that the DBEs play a role both in the biosynthesis of starch and in its degradation (Rahman et al., 1998; Dinges et al., 2003a; Dinges et al., 2003b). The metabolic interrelationships among the DBEs are further complicated by the fact that the isoamylase-type DBEs do not necessarily act independently of each other. Unlike the prokaryotic isoamylases, which do not have quaternary structure, the plant isoamylase-type DBEs exist as multimers (Dauvillée et al., 2001). In some instances the high molecular weight

complexes are homomultimers containing only ISA1, whereas in other instances ISA1 and ISA2 form a heteromultimeric complex. In dicot leaves and tubers the predominant form of isoamylase-type DBE is the ISA1/ISA2 heteromultimer (Hussain et al., 2003; Delatte et al., 2005), whereas in monocots it appears that ISA1 homomultimers predominate or that mixture of ISA1/ISA2 heteromultimers and ISA1 homomultimers may co exist (Fujita et al., 1999). The ISA2 component of the multimeric complex is likely to have a regulatory function as opposed to a direct catalytic role, because several amino acids that are nearly absolutely conserved in all members of the α-amylase-related superfamily, and are known in some instances to be involved in catalysis, are variant in all known ISA2 homologs. Furthermore, the unusual amino acid sequence in ISA2 is conserved across the plant kingdom, indicative of a conserved function (Hussain et al., 2003).

The existence of various quaternary structure forms of isoamylase-type DBEs, together with their functions in both degradation and synthesis of the glucan polymers of starch granules, gives rise to the hypothesis that different assembles can contribute to distinct physiological functions.

To further characterize the potential interactions between the three plant isoamylase-type DBE polypeptides, we investigated the tissue- and temporal specificity of gene expression for the *ISA1*, *ISA2*, and *ISA3* genes in Arabidopsis. The promoter of each gene was fused to the GUS reporter gene, and expression of each fusion protein was visualized in whole plant tissue. Specific pairs of genes are co-expressed as distinct sets, depending on the cell type and developmental stage. In particular, *ISA1* and *ISA2* are expressed together in several cell types and tissues in which *ISA3* expression is not detectable. Unexpectedly, *ISA2*- and *ISA3*- paired expression is observed in leaf hydathodes, trichomes, styles, and root tips although *ISA1* is not detectable. These data are consistent with the hypothesis that different quaternary structure assembles of distinct complexes involving ISA1, ISA2, and/or ISA3 mediate different physiological functions that are accomplished by the same α-(1→6) glucosidase activity. It also suggests the possibility that ISA2 functions in a complex with ISA3.

RESULTS

Informatics-based Co-expression of DBEs

In vitro studies of potato isoamylases (Hussain et al., 2003) and genetic analysis of Arabidopsis isoamylases mutants (Delatte et al., 2005) indicate that ISA1 and ISA2 may function as a dimer. An independent line of evidence consistent with the co-function of ISA1 and ISA2 could be the co-expression of these genes in specific cells, and also across particular developmental stages and environmental perturbations. Furthermore, such expression patterns might also provide information about the function of this complex. To get an idea of DBE expression across a variety of conditions, we evaluated Arabidopsis DBE RNA accumulation in a co-normalized set of public microarray data (Mentzen, 2006). RNA accumulation patterns were determined across 72 microarray experiments (over 1000 individual microarrays), representing a wide range of developmental conditions, and environmental and genetic perturbations using MetaOmGraph (MOG), a software developed to plot and analyze large sets of data (http://metnetdb.org/MetNet_MetaOmGraph.htm) (Figure 1a). Pearson correlation coefficients of RNA accumulation of the DBE genes were determined using MOG Affy.ath1.data1 data (Table 1). High correlation coefficient of RNA accumulation profile may reflect high correlation in function.

In these experiments, we evaluated each DBE gene independently for their RNA accumulation correlation with the 22746 genes represented on the Affymetrix Arabidopsis ATH1 chip. The microarray data analyses indicate a high co-expression of *ISA1* with *ISA2*. The RNA-accumulation Pearson correlation coefficient between *ISA1* (AT2G39930) and *ISA2* (AT1G03310) is 0.74; *ISA2* has the highest correlation coefficient with *ISA1*. *ISA1* and *ISA2* are co-expressed under certain conditions: their RNAs co-accumulate to high level in the first two true leaves (data not shown).

Surprisingly, *ISA2* and *ISA3* (AT4G09020) are also co-expressed. The RNA-accumulation Pearson correlation coefficient between *ISA2* and *ISA3* is 0.58. *ISA2* and *ISA3* are co-expressed under certain conditions: *ISA2* and *ISA3* RNAs co-accumulate to high level in siliques with seeds stage 4 and 5 (data not shown). Under most conditions, accumulation of *PU1* (AT5G04360) RNA is much lower than that of the *ISA* RNAs.

Interestingly, the microarray data analysis also indicates a high co-expression of *GWD1* and *GWD3* with the DBE genes, in particular with *ISA3* (Figure 1b). *GWD1* (At1g10760) and *GWD3* (At5g26570) water dikinases are involved in starch degradation (Yu et al., 2001;

Baunsgaard et al., 2005). These two plastidic enzymes catalyze starch phosphorylation by a dikinase-type mechanism in which the GWDs autophosphorylate via the β-phosphate of ATP, and this phosphate is transferred to either the C-6 or C-3 position of a starch glucosyl residue (Ritte et al., 2002; Mikkelsen et al., 2004; Kötting et al., 2004). Mutants of *GWD1* and *GWD3* are deficient in starch degradation (Yu et al., 2001; Baunsgaard et al., 2005). Because of GWD involvement in starch metabolism, and the information gathered from the microarray data studies indicating co-expression of the two Arabidopsis GWD genes with the DBE genes, we include *GWD1* and *GWD3* in Table 1.The RNA accumulation of *ISA3* is highly correlated with that of *GWD1* and *GWD3*, with RNA-accumulation Pearson correlation coefficients of 0.8 and 0.76, respectively. Out of the 22746 genes, these two GWD genes are among the top 5 that are highly co-expressed with *ISA3*.

Spatial Expression Patterns of DBE Genes

The co-expression of *ISA1* with *ISA2* and *ISA2* with *ISA3*, led us to investigate the expression of the DBE genes in more detail in cells. We designed constructs to express GFP/GUS reporters under the control of the *ISA1*, *ISA2*, *ISA3* or *PU1* promoter regions (Figure 2). In Arabidopsis, the promoter of a gene usually exists within about 1 kb upstream of the start codon. However, because the intergenic region for *ISA2* is only 888 bp (including a 347-bp intron located in the 5' UTR between the 46^{th} bp and the 47^{th} bp upstream of the ATG codon), we designed two promoter constructs for *ISA2*: ISA2p1, consisting of the intergenic region; and ISA2p2, consisting of the intergenic region plus part of upstream gene. Likewise, the intergenic sequence for *PU1* is only 296 bp, thus the three following promoter constructs were designed for *PU1*: PU1p1, consisting of the intergenic region; PU1p2, consisting of the intergenic region plus most of the upstream gene AT5G04350; and PU1p3, consisting of the intergenic region, plus the entire upstream gene AT5G04350, plus 669 bp of the upstream intergenic region between AT5G04350 and AT5G04347.

These DBE promoter::GUS constructs were transformed into Arabidopsis. For each of the seven constructs, at least ten independent transgenic lines in the T2 generation were screened for GUS expression (over 70 T2 transgenic lines in total). The lines were assessed at various stages of development from germination to senescence. Because the half-life of GUS

protein is 58 h and that of GFP is 26 h (Bio-Rad, Hercules, CA), to avoid possible fluctuations in the level of GFP/GUS accumulation associated with the diurnal cycle, plants were grown under continuous light for most developmental studies.

For the transgenic lines derived from the both *ISA2* promoter::GUS constructs, the patterns of GUS accumulation are indistinguishable; however, the level of GUS signal from the shorter ISA2p2 promoter is consistently less than that of the ISA2p1 promoter (data not shown). We conclude a functional promoter for ISA2 is within the upstream 883 bp sequence, and transgenic lines derived from this longer construct were used in the experiments described below. None of the three *PU1* promoter::GUS constructs (a total of 30 PU1 T2 transgenic lines were tested) directed detectable expression in the plant under any conditions studied, so the developmental function of these putative promoters could not be evaluated.

Expression of *ISA2* and *ISA3* first becomes evident at two days after imbibition (DAI) in hypocotyls and root tips (Figure 3a). *ISA1* expression is first detected at 3 DAI in the hypocotyl and cotyledon (data not shown). (*ISA2* and *ISA3* expression is also detected earlier than *ISA1* expression under a SD cycle of 8 h light/16 h dark. Under SD, seedling development is slowed, but the expression pattern is similar. *ISA2* and *ISA3* expression is first detected at 3 DAI in cotyledon, hypocotyls and root tips; there is no detection of *ISA1* expression at that time (data not shown)).

As the plants develop, *ISA1* and *ISA2* are co-expressed in mesophyll cells of cotyledons and emerging leaves (Figure 3b). As the seedlings grow, *ISA1* and *ISA2* are co-expressed in the shoot apex (Figure 3c); this pattern is maintained throughout the development. In addition, both *ISA1* and *ISA2* are co-expressed in the vasculature in all cotyledons (Figure 3b), rosettes (Figure 3e,f) and cauline leaves (data not shown) throughout expansion and maturation except that as leaves mature, *ISA1* expression in petiole and major veins is reduced (Figure 3d,f).

To better reveal the expression of ISA genes in the leaf vasculature and other tissues, semi-thin cross-sections were made from LR White-embedded GUS-stained samples (Figure 3g,h,k). *ISA1* and *ISA2* expression is localized within vascular bundles in all differentiating and differentiated cells of both xylem and phloem, and in bundle sheath cells (Figure 3g,h).

ISA1 has particularly strong expression in the bundle sheath cells (Figure 3g,h). No GUS staining is observed in the central vascular tissue of *ISA3* leaves (Figure 3g,h).

ISA1 expression occurs throughout the leaf epidermis (Figure 3g) and in guard cells of stomata (data not shown) including *ISA2* and *ISA3* are not detected in epidermis. In contrast, *ISA2* is expressed in whole petiole (Figure 3b), base of main vein (Figure 3e), and in the hydathodes (Figure 3e), throughout the development. But *ISA1* expression is conspicuously absent from hydathodes (Figure 3e,f), base of main vein and top of petiole (Figure 3d,f). *ISA3* expression is more limited, specifically to the edge of cotyledons and leaves: the hydathodes (Figure 3b,e,f), the associated vasculature of cotyledon and leaves at all stages of development and the leaf edge mesophyll cells (Figure 3b-f). Including co-expression in the hydathodes, *ISA2* and *ISA3* are both co-expressed in the leaf trichomes, where expression is not detected for *ISA1* (Figure 3i).

Because *ISA2* and *ISA3* are strongly co-expressed in hydathodes, and because little is known about the structure of Arabidopsis hydathodes, we examined DBE-hydathode associated expression in more detail. Apart from the passive release of water by transpiration, some plants can actively secrete water at openings or hydathodes, by a process called guttation (Esau, 1977; Fahn, 1979). Arabidopsis hydathode appears to have function mechanism to secret salt in guttation droplets (Pilot et al., 2004). Arabidopsis has active hydathode with vein ending and tightly packed epithem cells (sometime chloroplast-free, more or less modified mesophyll cells) surrounded by a sheath and one or more water pores (stomates that are usually permanently open) (Figure 3k); water passes through epithem cells and minerals may be reabsorbed (mediated by transfer cells) before water is forced out through the pore(s). *ISA2* and *ISA3* expression is high in all cell types of the hydathode (epithem) vasculature and in the mesophyll cells near the hydathode (Figure 3b,e,f,j,k). *ISA1* expression is not detected in any cell of this structure.

ISA1, *ISA2* and *ISA3* are all expressed in flower sepals and pedicels, and in stigma, throughout the maternal tissues of the ovules (integument) and ovary wall (Figure 3l,m). Cell-specific and development-specific co-expression of *ISA1* with *ISA2*, and also *ISA2* with *ISA3* extend to the flowers. *ISA1* and *ISA2* are co-expressed in the filament (Figure 3m) whereas *ISA2* and *ISA3* are co-expressed in the style (Figure 3m). Only *ISA2* expression is detectable in

the pollen (data not shown) and the receptacle (Figure 3l,m). No DBE expression is detectable in the petal or in the maternal-derived tissues of anther at any stage of development (Figure 3m).

The DBE gene expression patterns are also diverse with respect to siliques and ovules. In young developing seeds and embryos (1-3 DAF), only *ISA2* is highly expressed (Figure 3n). During silique development, *ISA1* is predominately expressed in the silique wall, *ISA2* is predominately in developing ovules, and *ISA3* expression decreases in the silique wall (Figure 3n,o).

ISA1 is expressed in the epidermis of roots at all stages of development (Figure 3p-r). *ISA2* and *ISA3* are co-expressed in the primary and lateral root tips (Figure 3p,q,s,t and data not shown), specifically in the root cap, columella and peripheral cap, to a lesser extent in the root meristem region, but not in the epidermis. Only *ISA2* is expressed in the root vasculature at the site of lateral root initiation, in the phloem, pericycle, endodermis and cortex (Figure 3p,t). As the lateral root develops, expression of *ISA2*, and to a lesser extent *ISA3*, becomes detectable throughout the root cortex and increases as the root matures (Figure 3r); *ISA2* only is also expressed in mature root vasculature (Figure 3r,t).

A summary of the DBE promoter::GUS expression profiling data is presented in Table S1 (Appendix D).

Sub-cellular Localization of ISA1 and ISA3 Proteins

Proteins can interact only when they are located in the same compartment. *Atisa1*, *Atisa2* and the *Atisa1*/*Atisa2* double mutant lack the same isoamylase activity band on a native PAGE gel of Arabidopsis chloroplast proteins (Zeeman et al., 1998a; Delatte et al., 2005). These data indicate that ISA1 and ISA2 are localized in the chloroplast, and that they form a protein complex. If ISA3 is also plastidic, it would be similarly theoretically possible for an ISA2-ISA3 interaction to occur. To get the location of ISA3, we developed 15 independent transgenic lines of Arabidopsis containing GFP/GUS fused to the N-terminal region of ISA3 (or ISA1), under the control of the *ISA3* (or *ISA1*) gene promoter region. The GFP fluorescence pattern in leaf confirms that both ISA3 and ISA1 are targeted to the plastids (Figure 4a). Plants transformed with a control plasmid containing the *ISA1* promoter but lacking the *ISA1*

targeting sequence show GFP signal dispersed throughout the cell (Figure 4b, left panel). Similar results were obtained with an ISA3 control plasmid (data not shown). Delatte and co-workers (Delatte et al., 2006) reported transient expression of AtISA3 protein in chloroplasts of Arabidopsis protoplasts in which AtISA3::GFP was transiently expressed.

Promoter Motif Analysis

DBE genes have complex spatial and developmental expression patterns. To better understand the factors associated with these patterns, we identified and compared putative *cis*-acting motifs present in the promoters of the DBEs. Since *GWD1* and *GWD3* are co-expressed with DBE genes at the transcription level (Figure 1), we also included their promoters in our analyses. Athena (http://www.bioinformatics2.wsu.edu/cgi-bin/Athena/cgi/visualize_select.pl; O'Connor et al., 2005) and Plant Care (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/) were used to identify the putative *cis*-acting motifs. Most of the motifs in the DBE and GWD gene promoters belong to one of four categories: light response, dehydration stress, cold stress or ABA signaling (Table 2).

All known Arabidopsis DBE and GWD genes contain light responsive motifs (Table 2), which might be related to their functions in the diurnal process of starch metabolism. *ISA1* contains eight light responsive motifs, *ISA2* contains three, *ISA3* contains seven, and *PU1*, *GWD1* and *GWD3* each contain six. *ISA3*, *PU1*, *GWD1* and *GWD3* each contain an Evening Element (EE) motif; this motif is necessary for dark-period-specific transcription (Michael and McClung, 2002). Harmer et al. (2000) found that of the 31 circadian genes in the Arabidopsis genome which contained the EE promoter motif, 30 peaked in expression at the end of light period of the next day. The DBE and GWD genes that contain the EE motif (*ISA3*, *PU1*, *GWD1* and *GWD3*) also have expression peaks at the end of light (Figure 5 for DBE; data not shown for GWD).

Each of these starch metabolic genes has cold responsive motifs in their promoters (Table 2). This might shed light on understanding the relation between starch metabolic genes and the stimulation of starch accumulation. The response of the plant to cold is in part mediated by ABA (Eckardt, 2001). The ABA signaling motifs found in all DBE and GWD genes might also be related to cold response. *ISA3* contains four cold responsive promoter

motifs (including "LTRE" (low temperature responsive element) and "DREB1A/CBF3" (dehydration-responsive element/C-repeat-binding, this motif interacts with "DREB1A/CBF3" transcription factor, which has induced expression in cold, and is important in freezing tolerance) (Maruyama et al., 2004) and six ABA signaling motifs. *GWD1* contains one cold responsive promoter motif and six ABA signaling motifs. *ISA1* contains seven ABA signaling motifs, *ISA2* contains one cold responsive motif and one ABA signaling motif, *PU1* contains one cold responsive motif and two ABA motifs, and *GWD3* contains two ABA motifs. *GWD1* contains the cold responsive promoter motif "CARrG". *GWD1* is reported to be important in the cold-induced development of cold tolerance in Arabidopsis (Yano et al., 2005).

DBE RNA Accumulation Profiles over a SD Diurnal Cycle

Because the promoter motif analysis showed the DBE genes contain light responsive promoter motifs and circadian motifs, we checked DBE RNA accumulation under a day/night cycle, to determine their expression pattern. Wild type (WT) Arabidopsis plants were grown under short day (SD, 8 h light/16 h dark) condition for six weeks. The 5th to 8th leaves were harvested at 11 time points. The same experiment was repeated once. RNAs from each time point of each replicate were isolated from the leaves and purified for Affymetrix Arabidopsis ATH1 chips. The microarray data were normalized by adjusting the mean expression level of each chip to 100.

The RNA accumulation profiles of DBE genes under short day conditions are shown in Figure 5(a). *ISA1* and *ISA2* are co-expressed over the diurnal cycle. Both have a big peak in the middle of the light and a smaller peak during the dark. *ISA2* maintains its high accumulation after the peak in the light till the darkness comes and decreases dramatically soon after, however *ISA1* decreases more gradually after its peak in the middle of the light. *ISA3* has only a single peak at the beginning of the dark phase. The overlap of high expressions of *ISA2* and *ISA3* near the beginning of dark is consistent with a possible co-function of *ISA2* and *ISA3*. The expression level of *PU1* doesn't change much over the diurnal cycle. It has a small peak at the beginning of the dark phase.

The co-expression of *ISA1* with *ISA2*, *ISA2* with *ISA3* during a diurnal cycle is also indicated by the RNA accumulation profiles from plants grown under a longer period of light

(Smith et al., 2004) (Figure 5b). *ISA1* and *ISA2* accumulation both peak at 4 h before the light comes in, which is the same as that of SD. *ISA3* increases at 4 h before dark and maintains the high expression until the dark comes in and reaches its peak at the beginning of the dark. *ISA3* shifts more toward light under 12 h light /12 h dark conditions.

Correlation of Expression of DBE Genes with Other Genes in the Arabidopsis Genome

Starch biosynthesis and degradation is a highly regulated process in which many enzymes and regulatory genes are involved. Genes with similar functions or in the same pathway might be highly correlated in RNA accumulation profiles. Thus, correlation analyses of transcript profiles could help identify genes which are functionally correlated. MOG was used to analyze the correlation of DBE expression with other genes in the Affy.ath1.data1 data.

Among the 22,746 genes on the Affymetrix Arabidopsis ATH1 chips, en toto**,** there are 30 genes which are highly correlated (correlation coefficient no less than 0.5) to all DBE genes (Appendix D, Table S2). Of these 30 genes, 12 are directly implicated in starch metabolism: a putative sugar transporter (AT4G04750), *BAM7* (AT4G00490), *ISA2*, *PU1*, *MEX1* (AT5G17520), *AAM3* (AT1G69830), *GWD1*, *GWD3*, *PHS1* (AT3G29320), *PHS2* (AT3G46970), *DPE1* (AT5G64860), and *DPE2* (AT2G40840). Except for the putative sugar transporter and *ISA2*, these genes are implicated to be involved in starch degradation (reviewed by Smith et al., 2005).

ISA3, *GWD1* and *GWD3* are specifically implicated to be involved in starch degradation (reviewed by Smith et al., 2005). In this public data set, 48 genes are co-expressed with all three of these genes (Appendix D, Table S3), including: 14 starch metabolic genes (*BE3* (AT2G36390), *BAM2* (AT2G32290), *ISA3*, *SEX4* (AT3G52180), *ISA2*, *PU1*, *MEX1*, *AAM3*, *GWD1*), *GWD3*, *PHS1*, *PHS2*, *DPE1*, and *DPE2*), 2 heat shock genes (an HSP (AT5G23240), and a small HSP (AT1G06460)), and 6 cold responsive genes (*COR314-TM2* (AT1G29390), *LTI6A* (AT4G30650), *COR15B* (AT2G42530), *KIN1* (AT5G15960), *COR15A* (AT2G42540), and *COR414-TM1* (AT1G29395)). 10 additional genes that are co-expressed are genes of unknown function, e.g., “expressed protein”.

Because the experimental data shows *GWD1* is induced in cold (Yano et al., 2005), promoter motif analysis shows DBE and GWD genes contain cold responsive motifs, and co-

expression patterns of the ISA3 and GWD genes with cold responsive genes, all indicate that starch metabolic genes could be cold responsive. We created a matrix of Pearson correlation coefficients among the starch metabolic genes and temperature responsive genes (Table 3). Some starch metabolic genes have high correlation coefficients between each other, like *ISA1* and *ISA2*; *DPE1*, *BE3* and *ISA3*; *BAM2* and *GWD3*. Other genes, which are indicated or reported to be involved in starch degradation (*SEX4*, *ISA3*, *MEX1*, *AAM3*, *PHS1*, *PHS2*, *GWD1*, *GWD3*, *DPE1* and *DPE2*) (Fordham-Skelton et al., 2002; reviewed by Smith et al., 2005), have high correlation coefficients between them.

Three experiments using cold treatments from the public data set were also examined. In these experiments, Arabidopsis plants were harvested after treated with cold at 5^{o}C for 3 h (Knight), at 4^{o}C for 24 h (Bramke), and at 4^{o}C for 10 days (Warren). Like the *COR314-TM2* (slow cold responsive) gene, *ISA3*, *GWD1* and *GWD3* expression increases at about one day after cold treatment, and is still maintained after ten days of cold treatment (Figure 6a). *ISA1*, *ISA2* and *PU1* do not respond to cold. Figure 6(b) shows the response of the 11 slow cold responsive genes that are mutually highly correlated.

Analyses of MOG indicate that of the three DREB family genes, two (*DREB1A*/*CBF3* (AT4G25480) and *DREB1C*/*CBF2* (At4g25470)) responded within 3 h to cold (Figure 6c). Their accumulation decreases after 24 h cold treatment, and is maintained at this level for at least ten days. Under the MOG experiments in the Affy.ath1.data1 dataset, cold is the only condition under which the DREB1A and DREB1C transcripts are highly accumulated (data not shown). Among the eleven slow responsive genes, nine are reported to be regulated by DREB1A/CBF3 (Maruyama et al., 2004): *COR314-TM2*, *COR15A*, *COR15B*, *KIN1* , *KIN2* , *LTI6A*, *COR78*, *COR47*, *LTI29*. Correlation analysis to *DREB1A*/*CBF3* in MOG showed no known starch metabolic genes are highly correlated with *DREB1A*/*CBF3* (its correlation coefficient is 0.12, 0.14, and 0.06 with *ISA3*, *GWD1* and *GWD3* respectively). We searched these cold responsive genes on TAIR: they are known to be involved in cold response; but their molecular functions are “unknown”. *ISA3* might be regulated by the transcription factor DREB1A/CBF3 and respond to cold similarly to the cold responsive genes that are regulated by DREB1A/CBF3 (Figure 6b); it is also possible that the slow cold responsive genes regulate *ISA3*’s response to cold.

Interestingly, all the cold responsive genes, which have high RNA-accumulation correlation coefficient with starch metabolic genes, are highly correlated to a gene: AT5G48250. AT5G48250 is a zinc finger (B-box type) family protein, belonging to C2C2-co-like family (which can promote flowering). It is highly co-expressed with *ISA3*, *GWD1* and *GWD3* (Figure 6d) not only under cold condition, but also across all the chips in the public data (data not shown). *AtGRP7/CCR2* (AT2G21660, *COLD CIRCADIAN RHYTHM AND RNA BINDING 2*) is another gene reported to be regulated by DREB1A/CBF3 (Maruyama et al., 2004). It is a glycine-rich RNA-binding protein that is involved in regulating cold circadian rhythm. It has similar accumulation profiles with *ISA3*, *GWD1* and *GWD3* not only under cold condition (Figure 6d), but also under other conditions (data not shown).

Totally there are 25 starch metabolic genes that are highly correlated (correlation coefficient 0.5 or more) with any of these 11 slow cold responsive genes indicating temperature regulates starch biosynthesis and degradation. Most of these genes are known to be involved in starch degradation. The expression of these starch metabolic genes and cold responsive genes are shown in Figure 6(e). Expression of some genes, in particular, *PHS2*, *BAM8*, *BE3*, *T6PP*, *BAM7*, *ISA3*, *SEX4*, *GWD1*, *SS1*, *GWD3*, *ADG2*, and *ADG1*, are increased by cold induction.

Cold and light responsive motifs exist commonly among the promoters of these 25 starch metabolic genes. The distributions of these promoter motifs in these starch metabolic genes (the genes with no obvious changes are excluded) are shown in Table S4 (Appendix D). Some genes, like *BAM7*, *ISA3*, *GWD1*, and *SS1*, contain more than five cold responsive promoter motifs. But *PHS2* and *BAM8* respond strongly to cold induction though they have only two cold responsive promoter motifs. It is very interesting that a starch synthesis involved gene *SS1*, with increased transcript after cold induction, contains 10 cold responsive promoter motifs. Including *ISA3*, *GWD1* and *GWD3*, the light responsive promoter motif EE also exists in the promoter regions of *PHS2*, *PGM1* and *PGM*.

Cellular Localization of DBE Genes under Changed Conditions

DBE genes have a lot of promoter motifs in their promoters. It indicates they might change their expression under different conditions. The expression of DBE genes under changed conditions was checked to test our promoter and correlation analyses.

To test the hypothesis that DBE genes might have different expressions under different light conditions or different conditions, we tried different experiments: leaves of plants grown under SD conditions till 15 DAI (two true leaves came out at this stage) were treated by coldness (4^oC) or dark. In contrast, none of the DBE genes responded transcriptionally to the bacteria infiltration or wounding (data not shown). No expression of PU1 was detected under any condition except very weak expression in root tip at one day after GA_3 application (data not shown).

Because some starch metabolic genes are up-regulated in response to cold, and because the cold response pattern is similar to some cold responsive genes that appear to be regulated by DREB1A, and because *ISA3* promoter contains a DREB1A/CBF3 cold responsive binding motif, we particularly evaluated the response of the DBE genes to cold. *ISA2* and *ISA3* are co-expressed in leaves after the plants become adapted to a cold environment. The plants almost cease growth after transferred to the cold under a SD diurnal cycle. The leaves retain small and become dark green. After 10 days in cold, *ISA1* and *ISA2* expressions are increased in cotyledon and reduced in leaves. *ISA3* expression is increased in cotyledon and leaf. After 41 days in cold, *ISA1* expression is still more in cotyledon than in leaf, *ISA3* and *ISA2* expression increases in the cold, whereas expression of *ISA1* decreases in leaves after prolonged cold (Figure 7). It is consistent with our in silico predictions that *ISA3* responds to cold and define this response as being particularly strong in leaf blade.

ISA3 responds more quickly to dark than *ISA1* and *ISA2*. By four days after transferred to dark, the expression of *ISA1* and *ISA2* was reduced in leaves. *ISA3* expression was decreased dramatically in cotyledon, leaf, and root. By 10 days in dark, *ISA1* expression was not detectable in expanding leaves, but was still visible in cotyledon. *ISA2* is still expressed in root. *ISA3* is even less expressed in root and cotyledon. By 18 days in the dark, the plants were pale and almost dying. *ISA1* and *ISA2* expression was minimal. *ISA3* had no detectable expression (Figure 7).

DISCUSSION

Starch DBEs hydrolyze α-(1→6) linkages and are critical for accumulation of normal levels, morphology, and composition of starch granules. They play a role in both starch biosynthesis and degradation (Rahman et al., 1998; Dinges et al., 2003a; Dinges et al., 2003b). Three models have been proposed for the functions of DBEs: a first model suggests the DBEs correctly modify the branching structure of starch (Ball et al., 1996; Myers et al., 2000); a second model suggests the DBEs degrade the phytoglycogen precursors (Ball et al., 1996; Zeeman, 1998b); a third model is that isoamylase affects starch granule initiation (Burton et al., 2002). But the functions and the relations among the members of the DBE gene family are not yet understood. This paper explores the relationship among the DBEs.

Both the temporal and spatial co-expression pattern of ISA-DBE genes is varied: a), *ISA1* and *ISA2* are co-expressed (e.g., leaf mesophyll cells and vasculatures, Figure 3d-f); b), *ISA2* and *ISA3* are co-expressed (e.g., leaf hydathodes and root tips, Figure 3j,s); c), all 3 isoamylase-type DBEs are co-expressed (e.g., leaf edge mesophyll cells and vasculatures, sepals, Figure 3b,m); d), *ISA1* and *ISA2*, but not *ISA3*, are expressed independently of other *ISA*s (e.g., *ISA1* in leaf and root epidermis and *ISA2* in the young embryo and at the site of lateral root initiation, Figure 3g,p).

The spatial patterns of co-expression of *ISA1* and *ISA2* during plant development are consistent with the evidence that ISA1 and ISA2 form a complex for starch synthesis (Hussain et al., 2003; Delatte et al., 2005). The co-expression of *ISA1* and *ISA2* both over the diurnal cycle (Figure 5), and also across a variety of conditions (Figure 1a and Table 1) support the co-function of ISA1 and ISA2. Because sequence analysis indicates that ISA2 is unlikely to be catalytic (Hussain et al., 2003), it was suggested that the ISA2 subunit recognizes the α-(1→6) branch point and helps the catalytic ISA1 subunit to remove the branch (Delatte et al., 2005). Our observation that ISA1 and ISA2 are highly accumulated in the cotyledons and the leaves (Figure 3b) indicates that ISA1 and ISA2 may be important for starch synthesis in these green tissues where photosynthesis and starch synthesis take place. The presence of ISA1 and ISA2 in all leaves and most tissues of the plant reflects the general importance of their function in starch synthesis.

The varied and specific patterns of co-expression of the *ISA2* and *ISA3* genes are somewhat surprising in light of the previous indication that ISA1 and ISA2 are involved in starch synthesis (Delatte et al., 2005), whereas ISA3 is important for degradation (Delatte et al., 2006). Despite the spatial-restricted expression pattern of *ISA3* (Figure 3b), *ISA2* and *ISA3* have co-expression as detected by analysis of spatial and temporal promoter::GUS expression Figure 3e,j,k). Also, *ISA2*/*ISA3* co-expression is evident in RNA accumulation profiles, not only over the diurnal cycle (Figure 5), but also across a variety of conditions (Figure 1a and Table 1). The co-expression of *ISA2* and *ISA3* opens the possibility that the ISA2 and ISA3 polypeptides might also form a protein complex. Both ISA3 protein (Figure 4a), and ISA2 protein (Zeeman et al., 1998a), are localized in the leaf chloroplast, also consistent with the possibility of an ISA2/ISA3 complex. Loss of ISA3 (*Atisa3* mutant) results in a starch excess phenotype (Delatte et al., 2006), which indicates ISA3 takes part in starch degradation. During the seed germination and post-germination, *ISA2* and *ISA3* expression is detected earlier than that of ISA1. One interpretation is that ISA2/ISA3 might be needed for starch degradation and/or possibly degradation of other forms of glucan to provide energy or material for seed germination and seedling growth. Thus, ISA2 could function in both starch biosynthesis and starch degradation. *Atisa2* contains less starch than WT (Delatte et al., 2005). This does not contradict our hypothesis that ISA2 could form a protein complex with ISA3 and takes part in starch degradation, as the biosynthetic process may dominate in the mutant, i.e., in a background of low starch as in *Atisa2*, a decrease in starch degradation might not be evident.

Auxin (IAA) has been reported to decrease accumulation of starch in tobacco (Miyazawa et al., 1999) and cotton (Bornman et al., 1966). It is interesting that *ISA2* and *ISA3* are co-expressed in the hydathode (Figure 3e,f,j). Hydathodes are a primary site for IAA synthesis (Roni and Cornelia, 2005), and IAA has been reported to reduce the expression of starch biosynthetic genes in cultured Bright Yellow-2 tobacco cells (Miyazawa et al., 1999). IAA correlates with pattern formation and orienting of cell division in Arabidopsis root development (Sabatini et al., 1999). The co-expression pattern of *ISA2* and *ISA3* in leaf and root (Figure 3b-f,p-s) is very similar to the expression of genes involved in IAA transport pathway: in leaf veins, hydathodes and trichomes, and in root cap cells (including columella cells) (Mattsson et al., 2003; Blilou et al., 2005). These evidences support that IAA might

regulate starch metabolism. Our data would suggest that ISA2, ISA3, GWD1 and GWD3 may participate in this IAA-induced starch accumulation.

Certain specific cells localized in hydathodes and root columella have no detectable *ISA1* expression but are with starch fluctuation. Hydathodes in Arabidopsis are water secreting glands (Pilot et al., 2004). The co-expression of *ISA2*/*ISA3* in the hydathodes suggests a possible regulatory role of water in starch accumulation. It has been reported that in root columella cells of Arabidopsis and radish, immediate degradation of starch, and a reduction in starch content take place in response to increased moisture (Takahashi et al., 2003). *ISA2* and *ISA3* are co-expressed in the columella cells of the root cap (Figure 3s,t) where there are starch grains in the amyloplasts that help to reorient the roots (Kiss, 2000) while *ISA1* is not detected. Taken together, these data suggest that ISA2 and ISA3 might be involved in root starch degradation in the root columella cells to control the root orientation and root growth for water in response to gravity (gravitropism) and water (hydrotropism). ISA2/ISA3 may play a role in the water/moisture modified starch metabolism.

All three isoamylases are highly expressed in flowers and young siliques (Figure 3l). They are expressed or co-expressed in many tissues of flower and silique. Flowering is a high energy demanding process. Carbohydrate plays an important role in floral transition and tissue differentiation (Corbesier et al., 1998). The high expression or co-expression of debranching enzymes in flower tissue indicates that starch hydrolysis takes important role in providing carbohydrate and energy in flower development and young embryo development.

Surprisingly, *ISA2* is detected in some tissues with no detectable *ISA1* or *ISA3*. ISA2 is predicted to be catalytically inactive (Hussain et al., 2003). A possible explanation could be the expression of other isoamylase gene is not at the detection level. The detection of *ISA2*-only in some tissues could also indicate that ISA2 might have a broader function other than forming a protein complex to determine amylopectin structure. ISA2 might function in conjunction or associate with factors other than debranching enzymes or ISA2 might function independently. Delatte et al. (2005) found there were different glucan structures in the *Atisa2* mutant in the chloroplasts from mesophyll cells and bundle sheath cells. The glucan structures included apparently normal starch granules, abnormal starch granules, and phytoglycogen. We have shown that ISA3 is expressed in some mesophyll cells and bundle sheath cells (Figure 3b,d-f).

This is consistent with and extends the suggestion by Delatte et al. (2005) that ISA3 has capacity for starch biosynthesis and can partially compensate for the loss of ISA1/ISA2 in the *Atisa2* mutant. The limited distribution of *ISA3* expression in only selected cell types and regions of the leaf (Figure 3b,d-f) could explain the partially WT/partially abnormal characteristic of the glucan structures observed in the *Atisa2* mutant.

Light responsive motifs commonly exist in the promoters of starch metabolic genes (Appendix D, Table S4). These motifs exist in DBE promoters (Table 2) would enable the DBEs to have diurnal or circadian rhythms. In particular, the EE promoter motif is necessary for dark-period-specific transcription (Michael and McClung, 2002). *ISA3*, *PU1* and other starch metabolic genes with EE promoter motif in the promoter region (*GWD1*, *GWD3*, *PHS2*, *PGM1*, and *DPE1*), all have peaks in RNA accumulation at the beginning of dark (Figure 5 and data not shown).

Promoters of starch metabolic genes are also enriched in cold responsive motifs (Appendix D, Table S4). *ISA3* contains a "DREB1A/CBF3" cold responsive promoter motif (Table 2). The DREB1A/CBF3 transcription factor takes part in activating other stress inducible genes' and is important for improving tolerance to stress (specifically freezing stress) by an increase in sucrose accumulation (Maruyama et al., 2004; Gilmour et al., 2000). *ISA3* which contains a DREB1A/CBF3 motif and also is increased in cold (Figure 7) might take part in starch degradation and provide sucrose under cold conditions to help to increase cold tolerance. The cold induced co-expression of both *ISA3* and *ISA2* (Figure 7) suggests *ISA3*, and *ISA2*, or other cold induced expression of starch metabolic genes (for example, *BAM8*, *BAM7*, *GWD1*, *GWD3*, *SEX4*, *PHS2*, *T6PP*, and *DPE2* (Appendix D, Table S4; Figure 6e); *GWD1* contains a "CARrG" cold responsive promoter motif and has been reported to be involved in the cold induced freezing tolerance development (Yano et al., 2005).) may also be implicated in cold stress.

The transcript accumulation patterns of *ISA3*, *GWD1* and *GWD3* are highly correlated to that of 14 starch metabolic genes. Of these starch metabolic genes, 11 (all except *BE3*, *BAM2* and *ISA2*) are reported to be involved in starch degradation (reviewed by Smith et al., 2005). Thus, the genes with similar profiles are also highly functionally correlated. The evidence that *ISA2* transcript profile is highly correlated to these starch degradative genes

(Table 3) is another indication that ISA2 may co-function with ISA3 in starch degradation. The other two genes, *BAM2* and *BE3*, with a predicted metabolic reaction but no clear physiological function, might also be involved in starch degradation.

This study defines cells of developmental stages and conditions in which the ISA genes may be important to starch metabolism. In addition, our data suggests that ISA2 and ISA3 might form a protein complex [ISA2/ISA3] and function together in starch metabolism, with ISA2 being a regulatory subunit in the complex. This hypothesis remains to be experimentally tested. Furthermore, the temporal and spatial co-expression of ISA1 and ISA2 are consistent with the idea that they are functioning together as a protein complex (ISA1/ISA2). Still more experiment analysis is needed to clarify their relations.

MATERIALS AND METHODS

Microarray Experiment

Plants used in this study were wild-type *Arabidopsis thaliana* (ecotype Columbia). Seeds were surface sterilized in 50% bleach/ 0.02% Triton X-100 solution for 7 min and rinsed in sterile distilled H_2O three times. The growth medium was buffered with 2.56 mM MES at pH 5.7 and contained 4.3 g/L Murashige and Skoog salts (GibcoBRL, Life Technologies, Rockville, MD), 1% sucrose, 1X B5 vitamins (100 μg/mL myo-inositol, 1 μg/mL pyridoxine hydrochloride, 1 μg/mL nicotinic acid, 10 μg/mL thiamine hydrochloride), and 0.6% (w/v) agargel (Sigma, St. Louis, MO). The seeds were placed in a growth chamber (73% RH at 22.2 ± 1.1°C) under a short day (SD) photoperiod (8 h/16 h) from fluorescent lamps (129 ± 16 $\mu M \cdot m^{-2} \cdot s^{-1}$ PAR). After 21 d, seedlings were transplanted to a sterile potting medium (LC1 Sunshine Mix, Sun Gro, Horticulture, Inc., Bellevue, WA) and fertilized with 1X Nutriculture Grower's Special Blend 21-8-18^{PLUS} (Plant Marvel Labs, Chicago, Il). Plants were arranged in randomized blocks (flats) using PROC PLAN of SAS (SAS Institute, Cary, NC) and returned to the same growth chamber and environmental conditions as used for germinating seeds. There are 21 pots in one flat with 2 plants / pot. Rosette leaves (#5 to #8) from 2 plants in one pot from each flat were harvested from six-week-old seedlings at 0, 0.5, 1, 4, 6, 8, 8.5, 9, 12, 14, 16, and 20 h during the SD photoperiod. Time points 0, 0.5, 1, 4, and 6 h were collected during the light cycle, and time points 8, 8.5, 9, 12, 14, 16, and 20 h were collected during the

dark cycle under a green light. Each sample consisted of leaves from sixteen plants. Samples were frozen in liquid N_2 immediately after harvest and stored at -80°C for RNA and protein extraction.

Plant Material

Plants (*Arabidopsis thaliana* ecotype Columbia) were grown in Sun Gro soil in 3" SQ pots and flats in a plant growth room under constant fluorescence white light at 22^{o}C.

Construction of DBE Promoter/Promoter+Target::GFP/GUS Fusion Vectors

Promoter::GFP/GUS fusion constructs were made for all four DBE genes by cloning amplified promoter region into a binary vector. In cases where the intergenic regions are short (no more than 1 kb), several constructs spanning various regions were made. A single construct was made for *ISA1*: ISA1p, a 1539-bp fragment upstream of the predicted ATG of *ISA1*, stops at 3' UTR of *ISA1*'s upstream gene. Primers used were ISA1ptarF (5'-GGTCACAAGAAAACAGAGAAC-3') and ISA1pR (5'-TGGAAGCTCGAAGAAGAG-3'). Two constructs were made for *ISA2*: ISA2p1 and ISA2p2 since the intergenic region is only 888 bp with a 347-bp intron within the 5' UTR of *ISA2*. For ISA2p1, an 883-bp fragment upstream of the predicted ATG of *ISA2* was selected. Primers used were ISA2p1F (5'-TAACAATGTTTAGGCTCCAA-3') and ISA2pR (5'-TGTTTAAACAAAGCACCAAA-3"). For ISA2p2, a 1456-bp fragment upstream of the predicted ATG of *ISA2* was used. Primers were ISA2p2F (5'-CCTTCACTATAAATTTCTTCTCAC-3') and ISA2pR (5'-TGTTTAAACAAAGCACCAAA-3"). A single construct was made for *ISA3*: ISA3p, a 1566-bp fragment upstream of the predicted ATG of *ISA3*, stops in the intergenic region. Primers used were ISA3ptarF (5'-GGATTTGGATTTCTGTATTT-3') and ISA3pR (5'-ATATCTTCGTTGCTTCAAC-3'). Three constructs were made for *PU1*: PU1p1, PU1p2 and PU1p3 since the intergenic region between *PU1* and its upstream gene is as short as 296 bp. PU1p1 is the 296-bp intergenic fragment upstream of the predicted ATG of *PU1*. Primers were PU1p1F (5'-TCTCTTCATATTTTCGATTGTTA-3') and PU1pR (5'-ATTCTTCTTCTTCTCCGATTG-3'). PU1p2 is a 729-bp fragment which stops at ATG of the upstream gene upstream of the predicted ATG of *PU1*. Primers were PU1p2F (5'-

AAATTAGAGTTTTATATTATCTTTG-3') and PU1pR (5'-ATTCTTCTTCTTCTCCGATTG-3'). PU1p3 is a-1406 bp fragment upstream of the predicted ATG of *PU1* and *PU1*'s upstream gene. Primers were PU1p3F (5'-AAGCGGATATTGGGCCTA-3') and PU1pR (5'-ATTCTTCTTCTTCTCCGATTG-3').

Promoter+target::GFP/GUS fusion constructs were made to express GFP/GUS fused to the N-terminal region of ISA1 (or ISA3), under the control of the *ISA1* (or *ISA3*) gene promoter region. ISA1 and ISA3 are predicted to be localized in the plastid by PSORT (http://psort.hgc.jp/form.html). The predicted plastid targeting peptides, based on PSORT, for the ISA1 and ISA3 proteins are 30 aa and 25 aa respectively; we therefore included the N-terminal 110 aa for ISA1, and the N-terminal 122 aa for ISA3 in the GFP expression constructs. We included longer sequence than the predicted plastid targeting peptides to make sure our constructs contained the plastid targeting peptides. Two constructs were designed for *ISA1* and *ISA3* respectively. ISA1ptar1 is a fragment from 1539 bp upstream of the predicted ATG of *ISA1* plus 320 bp of coding sequence (which includes 106 aa from ISA1). Primers were ISA1ptarF (5'-GGTCACAAGAAAACAGAGAAC-3') and ISA1tar1R (5'-TCGGAGAGAGAAATCAAGCAG-3'). ISA1ptar2 is a fragment from 1539 bp upstream of the predicted ATG of *ISA1* plus 330 bp of coding sequence (which includes 110 aa from ISA1). Primers used were ISA1ptarF (5'-GGTCACAAGAAAACAGAGAAC-3') and ISA1tar2R (5'-<u>GG</u>CTGACGGAGATCGGAGAG-3'. Marked <u>GG</u> are not from the genomic sequence. They were added to the beginning of 5' to make the sequence in frame). ISA3ptar1 is a 1566-bp fragment upstream of the predicted ATG of *ISA3* plus 101 bp (which includes 33 aa from ISA3). Primers were ISA3ptarF (5'-GGATTTGGATTTCTGTATTT-3') and ISA3tar1R (5'-ATGGGAATGGTGAAGCTACT-3'). ISA3ptar2 is a 1566-bp fragment upstream of the predicted ATG of *ISA3* plus 614 bp (three predicted introns were included. 368 bp from exon, which includes 122 aa from ISA3). Primers used were ISA3ptarF (5'-GGATTTGGATTTCTGTATTT-3') and ISA3tar2R (5'-CTCTGGGATAGTGATAAGCA-3').

Fragments including the promoter region or promoter plus target region were amplified by PCR from total DNA using the attB-adapted primers (5'-GGGGACAAGTTTGTACAAAAAAGCAGGCTTC-forward primer, 5'-GGGGACCACTTTGTACAAGAAAGCTGGGTC-reverse primer). After PCR-product

purification, the amplified promoter or promoter plus target region was cloned first into an entry vector pDONR221, then into a binary vector pBGWFS7 (Karimi et al., 2002; Flanders Interuniversity Institute for Biotechnology, 2003) upstream of promoterless GFP/GUS gene (Gateway™ Cloning Technology, Invitrogen, CA, 2003). Gateway vectors were transformed into *Agrobacterium tumefaciens* strain GV3101.

Arabidopsis Transformation and Selection

Arabidopsis plants were transformed using the floral dip method (Clough and Bent, 1998). Transformed Arabidopsis plants were selected based on Bar resistance conferred by the T-DNA. T2 seeds from at least 15 independently-transformed lines for each construct were harvested for GUS and GFP screening.

Histochemistry and Microscopy

Transgenic T2 seedlings were germinated in soil in pots. For transgenic lines containing promoter::GFP/GUS constructs, plants were harvested at various stages of development. A minimum of 3 plants were harvested for each independently-transformed line; plants or organs from the same line were stained in the same tube. Organs were stained in GUS-staining solution (Triton/Ethanol stock (Triton X-100: Ethanol: Water, 1:4:5): 0.5M PH7.0 KPO4 Buffer: 10mg/ml X-gluc in dimethyl sulfoxide: 0.1M PH7.0 Potassium ferricyanide: 0.1M PH7.0 Potassium ferrocyanide, 5:470:25:2:2) at room temperature overnight or for 3 h (Jefferson, 1987**)**. After destained in 70% ETOH, organs were fixed overnight in FAA (50% EtOH, 5% acetic acid, 3.7% formaldehyde) and cleared in Herr's buffer, composed of 85% lactic acid, chloral hydrate, phenol, clove oil, and xylene (2:2:2:2:1), under vacuum for 3 weeks (Herr, 1971). Staining patterns were observed and documented using an Olympus stereomicroscope in the Bessey Microscopy Facility (Iowa State University, Ames, IA).

Cytological Techniques and Microscopy

GUS-stained tissue blocks were dehydrated in a graded ethanol series from 30% to 95% and then in 100% ethanol for 2 h at room temperature. A gradual infiltration used ethanol/LR white mixtures of 2/1, 1/1, and 1/2 for 2 h each at room temperature. A final infusion with

100% LR white was performed overnight at 4^{o}C and embedded in an acrylic resin (LR white, London Resin Company, London, UK). Polymerization was done overnight in gelatin capsules at 60^{o}C. After hardening, the specimens were cut into two m sections using glass knives fitted to a Reichert microtome Ultracut. Semi-thin sections (2 m) were stained with 2.0% (w/v) Safranin O for 20 second. After staining, the sections were dehydrated through an ethanol series to 100% ethanol, placed in xylene, and then covered with mounting medium (Permount, Fisher) and a cover slip. Slide-mounted sections were viewed using a Zeiss Axioplan compound microscope with a Zeiss AxioCamHRc digital camera (Carl Zeiss, Inc., Thornwood, NY).

GFP and Confocal Laser Scanning Microscopy

For promoter+target constructs, young leaves (leaves three to six) from 15 DAI (days after planting) plants were placed in water on slides, covered with cover glass. Slides were observed and documented under a Leica TCS NT laser scanning microscope system in the Confocal Microscopy Facility (Iowa State University, Ames, IA). Under confocal microscope, GFP signal shows green color and the auto fluorescence shows red color. The AR/KR laser and 488 nm/568 nm FITC/TRITC wavelengths were used. We found that yellowish or damaged or dying part of the leaf would show green color under confocal microscope. Only the healthy, green leaves were selected to detect real GFP signal. Both microscope digital images were processed in Adobe Photoshop 7.0 (Adobe, San Jose, CA).

Transcriptomics Analysis

MetaOmGraph (http://www.metnetdb.org/MetNet_MetaOmGraph.htm) was used to analyze expression patterns of DBE and starch related genes. The experimental data and metadata from 70 experiments comprising 956 Affymetrix ATH1 microarray slides were obtained from two online microarray depositories: NASCArrays (http://affymetrix.arabidopsis.info/narrays/experimentbrowse.pl, Craigon et al., 2004) and PLEXdb (http://www.plexdb.org/, Shen et al., 2005). The data were normalized to the same range by scale normalization. The biological replicates were averaged to yield 424 samples.

This normalized data is available online (http://www.metnetdb.org/MetNet_MetaOmGraph.htm).

ACKNOWLEDGMENTS

We thank Dr. Jack Horner, Tracey Pepper and Randall Den Adel in Bessey Microscope Facility and Margie Carter in Confocal Microscopy Facility for their help in using the equipment.

LITERATURE CITED

Ball S, Guan H-P, James M, Myers A, Keeling P, Mouille G, Buleon A, Colonna P, and Preiss J (1996) From glycogen to amylopectin: A model for the biogenesis of the plant starch granule. Cell **86:** 349–352

Baunsgaard L, Lutken H, Mikkelsen R, Glaring MA, Pham TT, Blennow A (2005) A novel isoform of glucan, water dikinase phosphorylates pre-phosphorylated alpha-glucans and is involved in starch degradation in Arabidopsis. Plant J **41:** 595–605

Beatty MK, Rahman A, Cao H, Woodman W, Lee M, Myers AM, James MG (1999) Purification and molecular genetic characterization of ZPU1, a pullulanase-type starch-debranching enzyme from maize. Plant Physiol **119:** 255–266

Blilou I, Xu J, Wildwater M, Willemsen V, Paponov I, Friml J, Heidstra R, Aida M, Palme K and Scheres B (2005) The PIN auxin efflux facilitator network controls growth and patterning in Arabidopsis roots. Nature **433:** 39–44

Bornman CH, **Addicott FT**, **Spurr AR** (1966) Auxin and gibberellin effects on cell growth and starch during abscission in cotton. Plant Physiol **41:** 871–876

Burton RA, Jenner H, Carrangis L, Fahy B, Fincher GB, Hylton C, Laurie DA, Parker M, Waite D, van Wegen S, Verhoeven T, Denyer K (2002) Starch granule initiation and growth are altered in barley mutants that lack isoamylase activity. Plant J **31:** 97–112

Burton RA, Zhang XQ, Hrmova M, Fincher GB (1999) A single limit dextrinase gene is expressed both in the developing endosperm and in germinated grains of barley. Plant Physiol **119:** 859–871

Bustos R, Fahy B, Hylton CM, Seale R, Nebane NM, Edwards A, Martin C, Smith AM (2004) Starch granule initiation is controlled by a heteromultimeric isoamylase in potato tubers. Proc Natl Acad Sci USA **101:** 2215–2220

Clough SJ, Bent AF (1998) Floral dip: a simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana*. Plant J **16:** 735–743

Corbesier L, Lejeune P, Bernier G (1998) The role of carbohydrates in the induction of flowering in *Arabidopsis thaliana*: comparison between the wild type and a starchless mutant. Planta **206:** 131–137

Craigon DJ, James N, Okyere J, Higgins J, Jotham J, May S (2004) NASCArrays: a repository for microarray data generated by NASC's transcriptomics service. Nucleic Acids Res **32:** D575–D577

Dauvillée D, Colleoni C, Mouille G, Morell MK, D'Hulst C, Wattebled F, Liénard L, Delvallé D, Ral J-P, Myers AM, et al (2001) Biochemical characterization of wild type and mutant isoamylases of *Chlamydomonas reinhardtii* supports a function of the multimeric enzyme organization in amylopectin maturation. Plant Physiol **125:** 1723–1731

Delatte T, Trevisan M, Parker ML, Zeeman SC (2005) Arabidopsis mutants Atisa1 and Atisa2 have identical phenotypes and lack the same multimeric isoamylase, which influences the branch point distribution of amylopectin during starch synthesis. Plant J **41:** 815–830

Delatte T, Umhang M, Trevisan M, Eicke S, Thorneycroft D, Smith SM, Zeeman SC (2006) Evidence for distinct mechanisms of starch granule breakdown in plants. J Biol Chem **281:** 12050–12059

Dinges JR, Colleoni C, James MG, Myers AM (2003a) Mutational analysis of the pullulanase-type debranching enzyme of maize indicates multiple functions in starch metabolism. Plant Cell **15:** 666–680

Dinges JR, Colleoni C, Myers AM, James MG (2001) Molecular structure of three mutations at the maize *sugary1* locus and their allele-specific phenotypic effects. Plant Physiol **125:** 1406–1418

Dinges JR, James MG, Myers AM (2003b) Genetic analysis indicates maize pullulanase- and isoamylase-type starch debranching enzymes have partially overlapping functions in starch metabolism. J Appl Glycosci **50:** 191–195

Doehlert DC, Knutson CA (1991) Two classes of starch debranching enzymes from developing maize kernels. J Plant Physiol **138:** 566–572

Eckardt NA (2001) Luc genetic screen illuminates stress-responsive gene regulation. Plant Cell **13:** 1969–1972

Esau K (1977) Anatomy of seed plants. John Wiley and Sons, New York

Fahn A (1979) Plant Anatomy, 4th Edition. Pergamon Press, New York

Fordham-Skelton AP, Chilley P, Lumbreras V, Reignoux S, Fenton TR, Dahm CC, Pages M, Gatehouse JA (2002) A novel higher plant protein tyrosine phosphatase interacts with SNF1-related protein kinases via a KIS (kinase interaction sequence) domain. Plant J **29:** 705–715

Fujita N, Kubo A, Francisco PB Jr, Nakakita M, Harada K, Minaka N, Nakamura Y (1999) Purification, characterization, and cDNA structure of isoamylase from developing endosperm of rice. Planta **208:** 283–293

Gilmour SJ, Sebolt AM, Salazar MP, Everard JD, Thomashow MF (2000) Overexpression of the Arabidopsis *CBF3* transcriptional activator mimics multiple biochemical changes associated with cold acclimation. Plant Physiol **124:** 1854–1865

Harmer SL, Hogenesch JB, Straume M, Chang HS, Han B, Zhu T, Wang X, Kreps JA, Kay SA (2000) Orchestrated transcription of key pathways in Arabidopsis by the circadian clock. Science **290:** 2110–2113

Herr JM (1971) A new clearing-squash technique for the study of ovule development in angiosperms. Am J Bot **58:**785–790

Hussain H, Mant A, Seale R, Zeeman S, Hinchliffe E, Edwards A, Hylton C, Bornemann S, Smith AM, Martin C, Bustos R (2003) Three isoforms of isoamylase contribute different catalytic properties for the debranching of potato glucans. Plant Cell **15:** 133–149

Invitrogen (2003) Gateway™ Cloning Technology

James MG, Robertson DS, Myers AM (1995) Characterization of the maize gene *sugary1*, a determinant of starch composition in kernels. Plant Cell **7:** 417–429

Jefferson RA (1987) Assaying chimeric genes in plants: The GUS gene fusion system. Plant Mol Biol Rep **5:** 387–405

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

Kiss JZ (2000) Mechanism of the early phases of plant gravitropism. Crit Rev Plant Sci **19:** 551–573

Kötting O, Pusch K, Tiessen A, Geigenberger P, Steup M, Ritte G (2005) Identification of a novel enzyme required for starch metabolism in Arabidopsis leaves. The phosphoglucan, water dikinase. Plant Physiol **137:** 242–252

Kubo A, Fujita N, Harada K, Matsuda T, Satoh H, Nakamura Y (1999) The starch-debranching enzymes isoamylase and pullulanase are both involved in amylopectin biosynthesis in rice endosperm. Plant Physiol **121:** 399–409

Manners DJ (1997) Observations on the specificity and nomenclature of starch debranching enzymes. J Appl Glycosci **44:** 83–85

Maruyama K, Sakuma Y, Kasuga M, Ito Y, Seki M, Goda H, Shimada Y, Yoshida S, Shinozaki K, Yamaguchi-Shinozaki K (2004) Identification of cold-inducible downstream genes of the Arabidopsis DREB1A/CBF3 transcriptional factor using two microarray systems. Plant J **38:** 982–993

Mattsson J, Ckurshumova W, Berleth T (2003) Auxin signaling in Arabidopsis leaf vascular development. Plant Physiol **131:** 1327–1339

Mentzen IW (2006) From pathway to regulon in Arabidopsis. Ph.D. dissertation. Iowa State University, Ames, IA

Michael TP, McClung CR (2002) Phase-specific circadian clock regulatory elements in *Arabidopsis thaliana*. Plant Physiol **130:** 627–638

Mikkelsen R, Baunsgaard L, Blennow A (2004) Functional characterization of alpha-glucan,water dikinase, the starch phosphorylating enzyme. Biochem J **377:** 525–532

Miyazawa Y, Sakai A, Miyagishima S, Takano H, Kawano S, Kuroiwa T (1999) Auxin and cytokinin have opposite effects on amyloplast development and the expression of starch synthesis genes in cultured bright yellow-2 tobacco cells. Plant Physiol **121:** 461–469

Myers AM, Morell MK, James MG, Ball SG (2000) Recent progress toward understanding biosynthesis of the amylopectin crystal. Plant Physiol **122:** 989–997

O'Connor TR, Dyreson C, Wyrick JJ (2005) Athena: a resource for rapid visualization and systematic analysis of Arabidopsis promoter sequences. Bioinformatics. In press

Pilot G, Stransky H, Bushey DF, Pratelli R, Ludewig U, Wingate VP, Frommer WB (2004) Overexpression of GLUTAMINE DUMPER1 leads to hypersecretion of glutamine from hydathodes of Arabidopsis leaves. Plant Cell **16:** 1827–1840

Rahman A, Wong K, Jane J, Myers AM, James MG (1998) Characterization of SU1 isoamylase, a determinant of storage starch structure in maize. Plant Physiol **117:** 425–435

Ritte G, Lloyd JR, Eckermann N, Rottmann A, Kossmann J, Steup M (2002) The starch-related R1 protein is an alpha-glucan, water dikinase. Proc Natl Acad Sci USA **99:** 7166–7171

Roni A, Cornelia IU (2005) Free auxin production and vascular differentiation in Arabidopsis leaves. Plant Physiology Online, a companion to Plant Physiology, Third Edition by Lincoln Taiz and Eduardo Zeiger, published by Sinauer Associates

Sabatini S, Beis D, Wolkenfelt H, Murfett J, Guilfoyle T, Malamy J, Benfey P, Leyser O, Bechtold N, Weisbeek P, Scheres B (1999) An auxin-dependent distal organizer of pattern and polarity in the Arabidopsis root. Cell **99:** 463–472

Shen L, Gong J, Caldo RA, Nettleton D, Cook D, Wise RP, Dickerson JA (2005) BarleyBase—an expression profiling database for plant genomics. Nucleic Acids Res **33:** D614–D618

Smith AM, Zeeman SC, Smith SM (2005) Starch degradation**.** Annu Rev Plant Biol **56:** 73–98

Smith SM, Fulton DC, Chia T, Thorneycroft D, Chapple A, Dunstan H, Hylton C, Zeeman SC, Smith AM (2004) Diurnal changes in the transcriptome encoding enzymes of starch metabolism provide evidence for both transcriptional and posttranscriptional regulation of starch metabolism in Arabidopsis leaves. Plant Physiol **136:** 2687–2699

Takahashi N, Yamazaki Y, Kobayashi A, Higashitani A, Takahashi H (2003) Hydrotropism interacts with gravitropism by degrading amyloplasts in seedling roots of Arabidopsis and radish. Plant Physiol **132:** 805–810

Wattebled F, Dond Y, Dumez S, Delvalle D, Planchot R, Berbezy P, Vyas D, Colonna P, Chatterjee M, Ball S, and D'Hulst C (2005) Mutants of Arabidopsis lacking a chloroplastic

isoamylase accumulate phytoglycogen and an abnormal form of amylopectin. Plant Physiol **138:** 184–195

Yano R, Nakamura M, Yoneyama T, Nishida I (2005) Starch-related alpha-glucan/water dikinase is involved in the cold-induced development of freezing tolerance in Arabidopsis. Plant Physiol **138:** 837–846

Yu TS, Kofler H, Hausler RE, Hille D, Flugge UI, Zeeman SC, Smith AM, Kossmann J, Lloyd J, Ritte G, Steup M, Lue WL, Chen J, Weber A (2001) The Arabidopsis *sex1* mutant is defective in the R1 protein, a general regulator of starch degradation in plants, and not in the chloroplast hexose transporter. Plant Cell **13:** 1907–1918

Zeeman SC, Northrop F, Smith AM, ap Rees T (1998a) A starch-accumulating mutant of *Arabidopsis thaliana* deficient in a chloroplastic starch-hydrolysing enzyme. Plant J **15:** 357–365

Zeeman SC, Umemoto T, Lue WL, Au-Yeung P, Martin C, Smith AM, Chen J (1998b) A mutant of Arabidopsis lacking a chloroplastic isoamylase accumulates both starch and phytoglycogen. Plant Cell **10:** 1699–1712

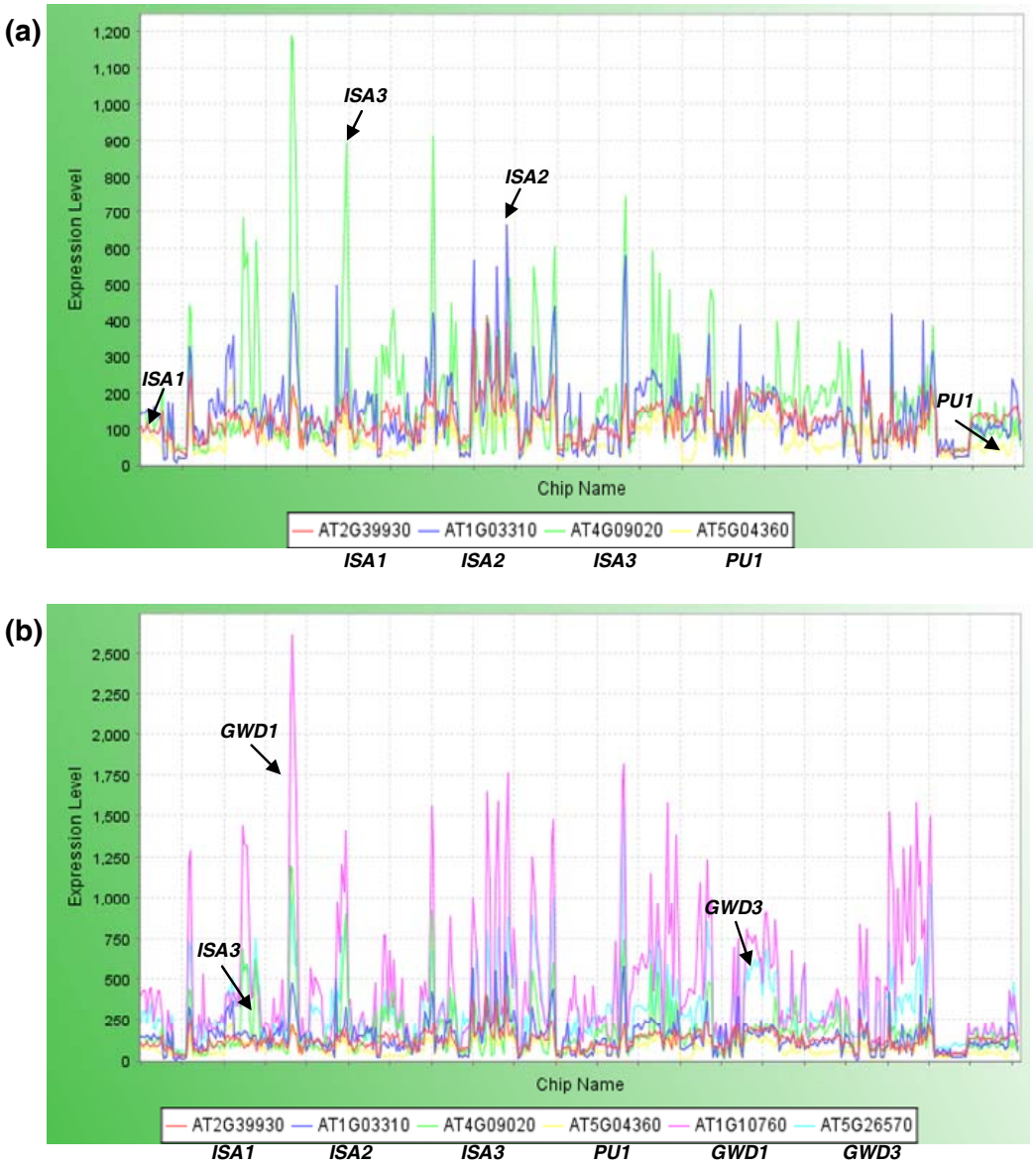


Figure 1. RNA accumulation profiles of starch metabolic genes.
RNA accumulation of DBE genes (a), of DBE genes plus *GWD1* and *GWD3* (b). The average expression level for each chip is 100.

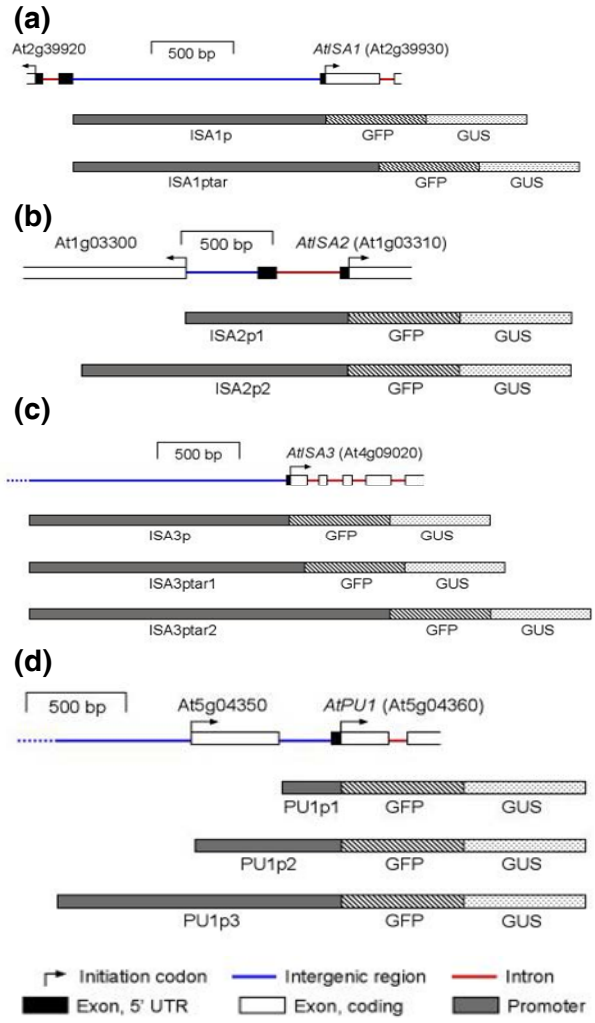


Figure 2. Promoter/promoter+target::GFP/GUS constructs for DBE genes.
(a) Constructs for *ISA1*, At2g39930.
(b) Constructs for *ISA2*, At1g01130.
(c) Constructs for *ISA3*, At4g09020.
(d) Constructs for *PU1*, At5g04360.

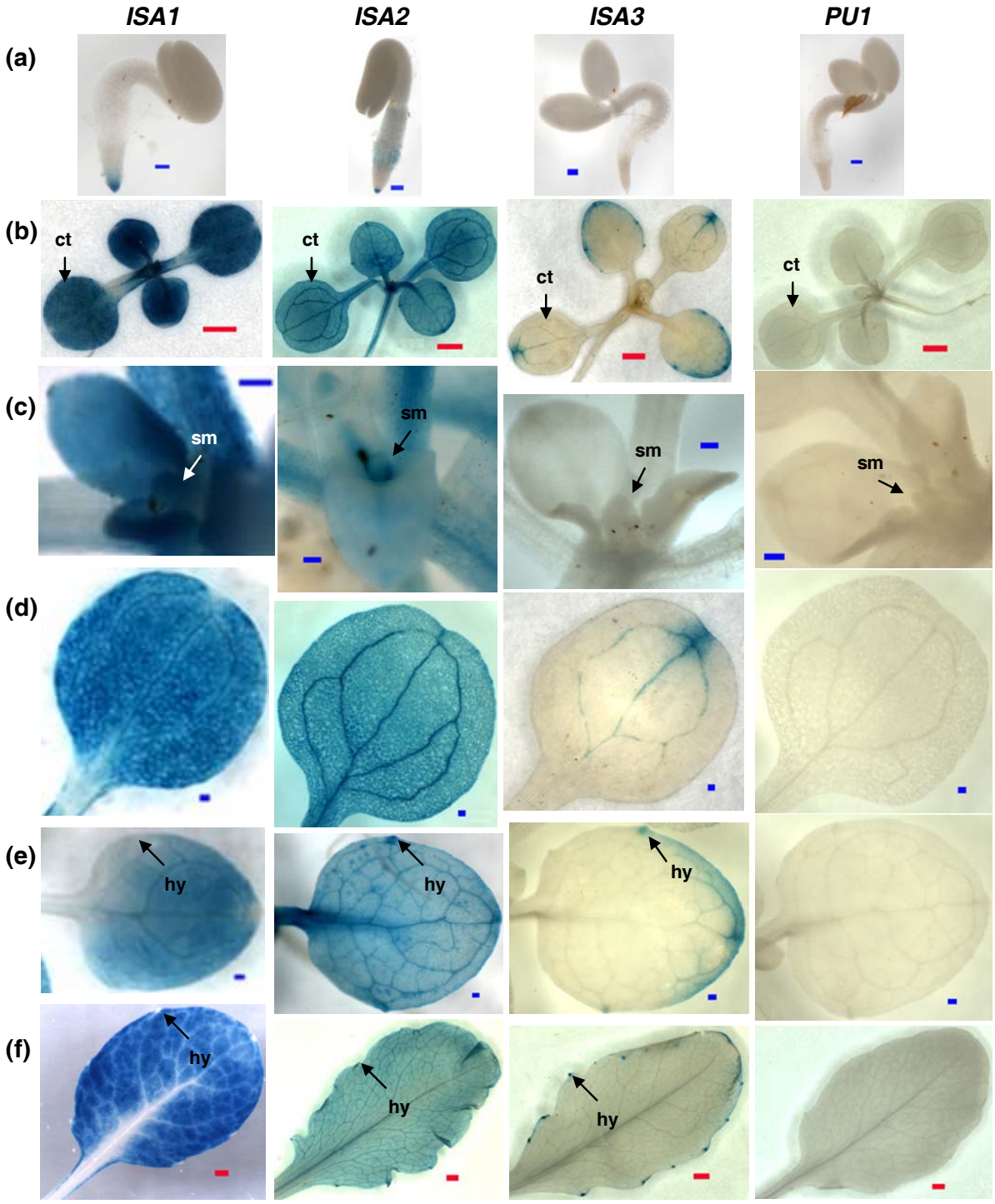


Figure 3.

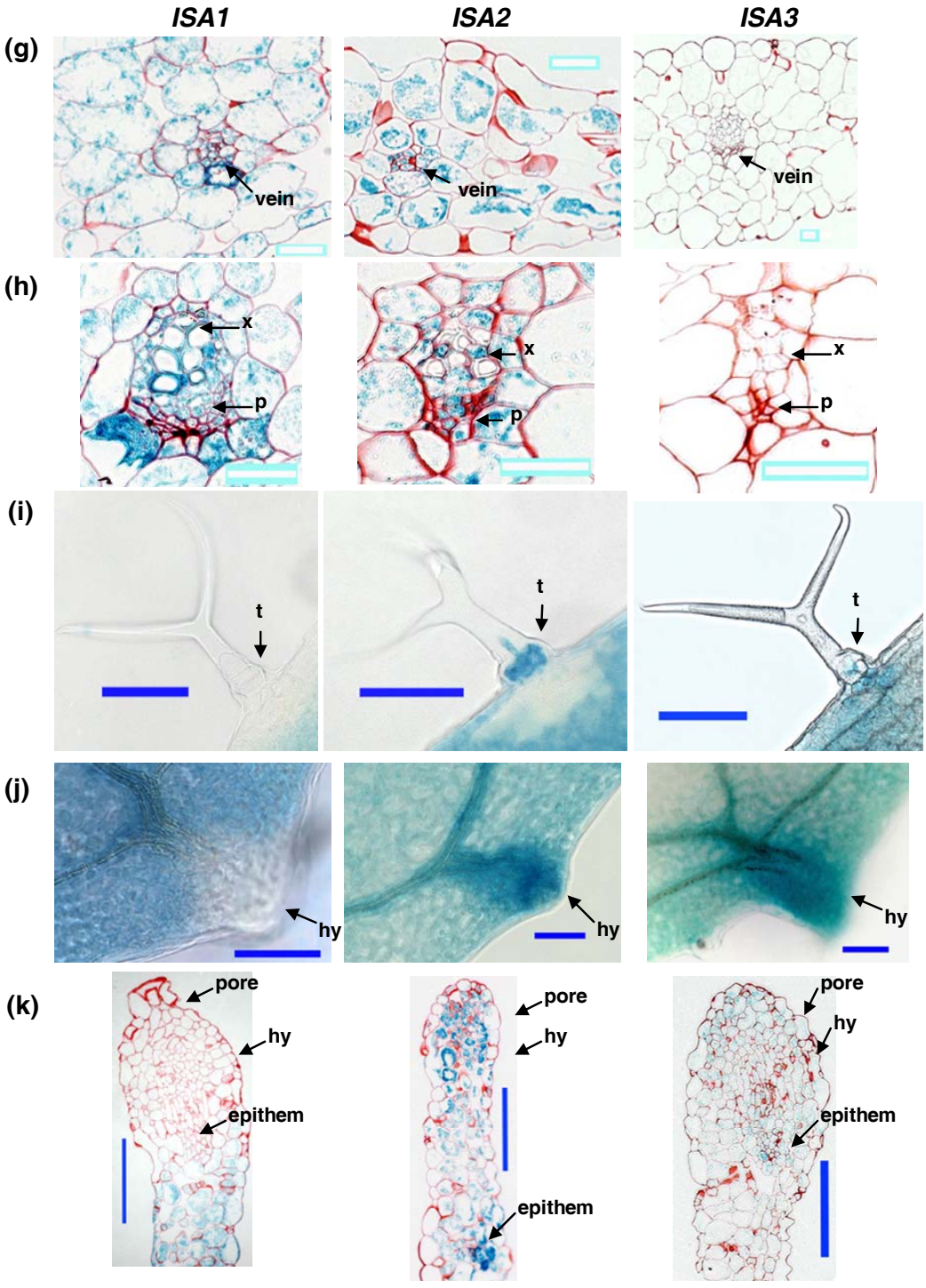


Figure 3. (continued).

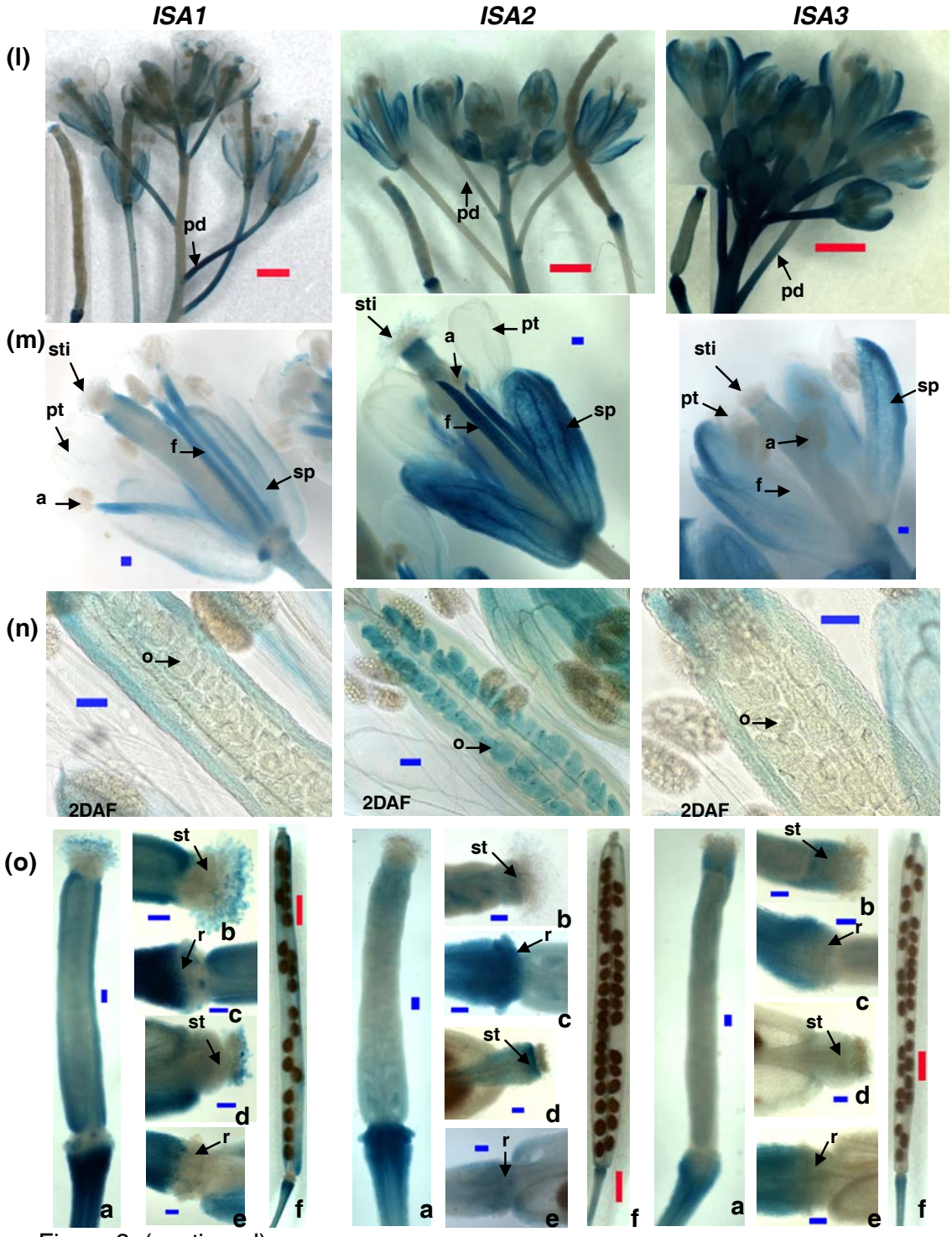


Figure 3. (continued).

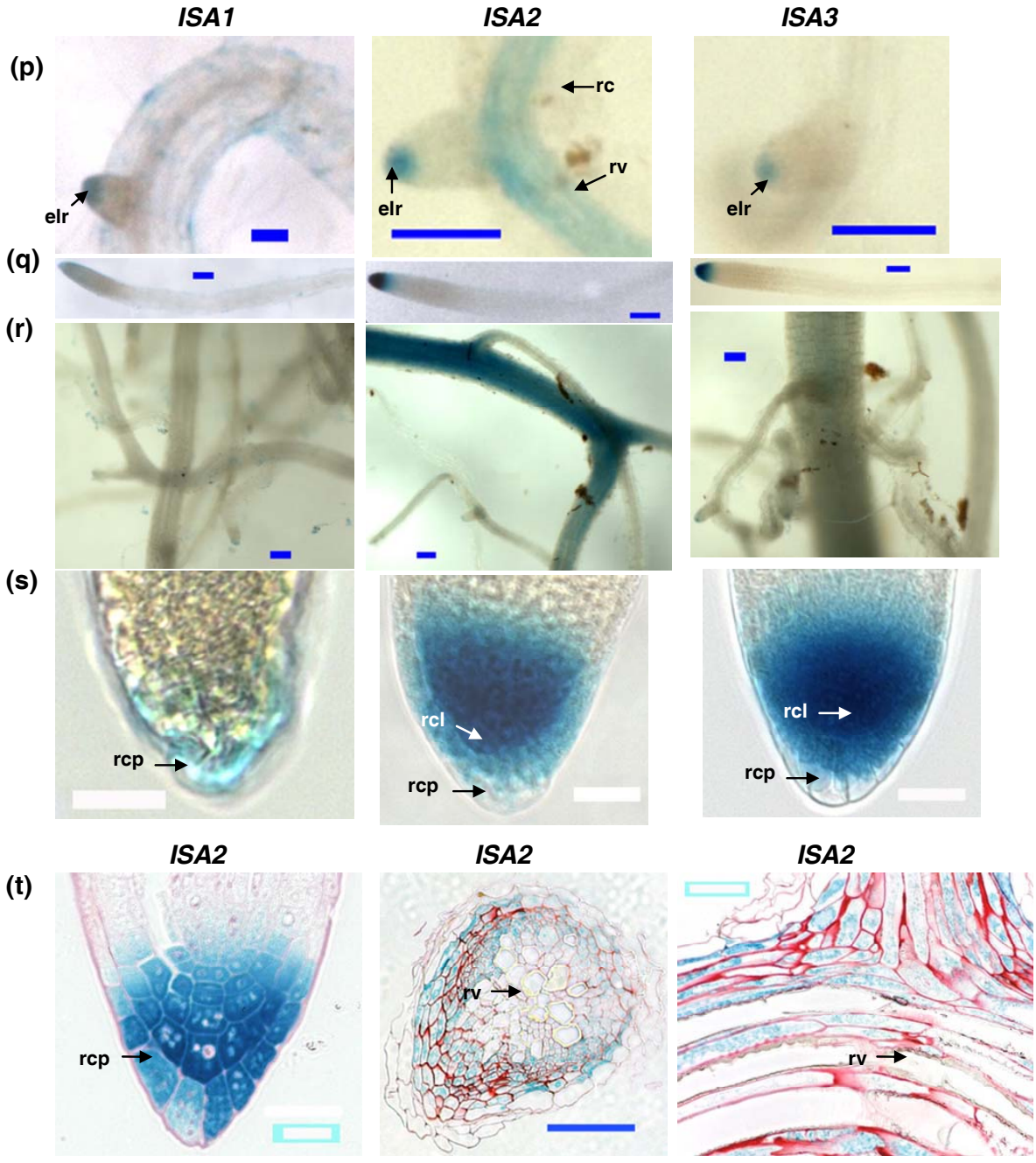


Figure 3. (continued).

Figure 3. (continued). DBE genes are differently expressed in Arabidopsis over different developmental stages.
Spatial expression patterns of DBE genes promoter::GUS. Sections are counterstained with safranin O to visualize cells. *ISA1* and *ISA2* are co-expressed in leaf mesophyll, vasculature, shoot meristem, filament, and part of leaf petiole. *ISA2* and *ISA3* are co-expressed in hydathode, root tip, root meristem, style, trichome base, part of root vasculature. DBE expression in germinating seedling (a), seedling (b), shoot meristem (c), cotyledon (d), young leaf (e), old leaf (f), leaf cross section (g-h), trichome (i), hydathode (j), hydathode cross section (k), flowers and buds (l), flower (m), embryos in young silique (n), siliques (o), and roots (p-t). a=anther, ct=cotyledon, elr=emerging lateral root, f=filament, h=hydathode, o=ovule, p=phloem, pd=pedicle, pt=petal, r=receptacle, rc=root cortex, rcl=root collumella, rcp=root cap, rt=root tip, rv=root vasculature, sm=shoot meristem, sp=sepal, sti=stigma, st=style, t=trichome, x=xylem. Red bar=1mm, blue bar=0.1mm, white bar=20µm.

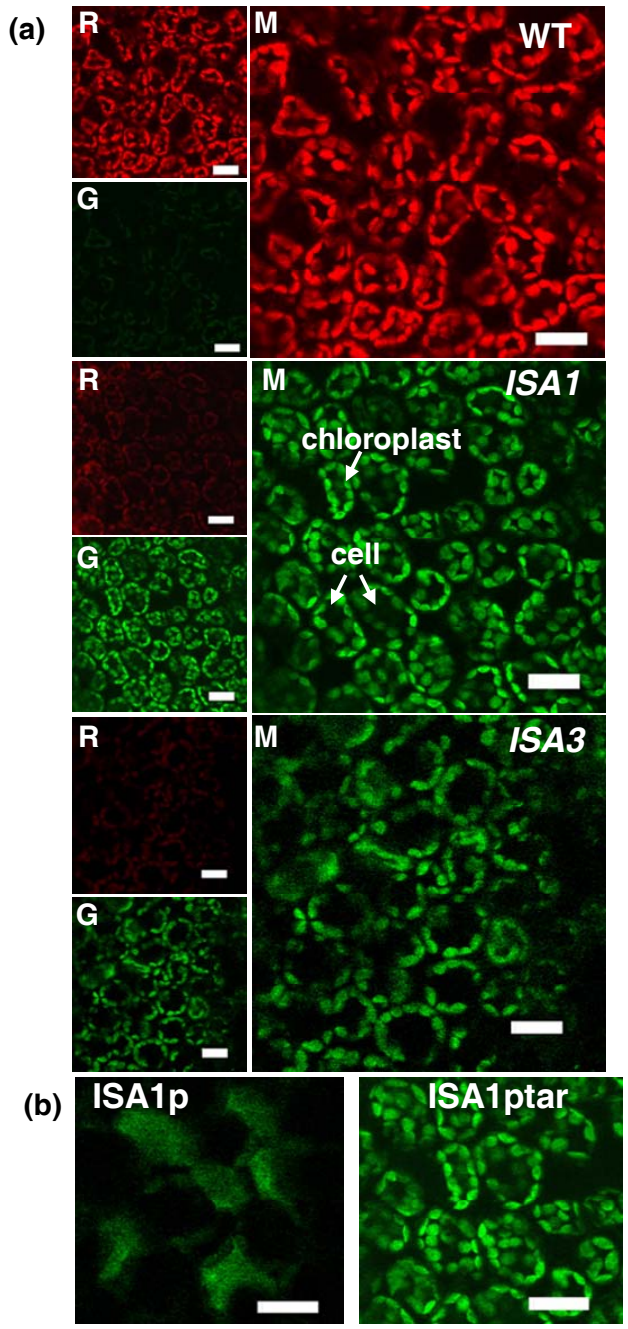


Figure 4. ISA1 and ISA3 are localized in the plastids.
(a) *ISA1* and *ISA3* promoter-target sequence-GFP fusion in leaf.
(b) Signal comparison of promoter::GFP and promoter+target::GFP for *ISA1*.
White bar=20μm. "R" means autofluorescence signal; "G" means GFP signal; "M" means merged image from autofluorescece and GFP signal images.

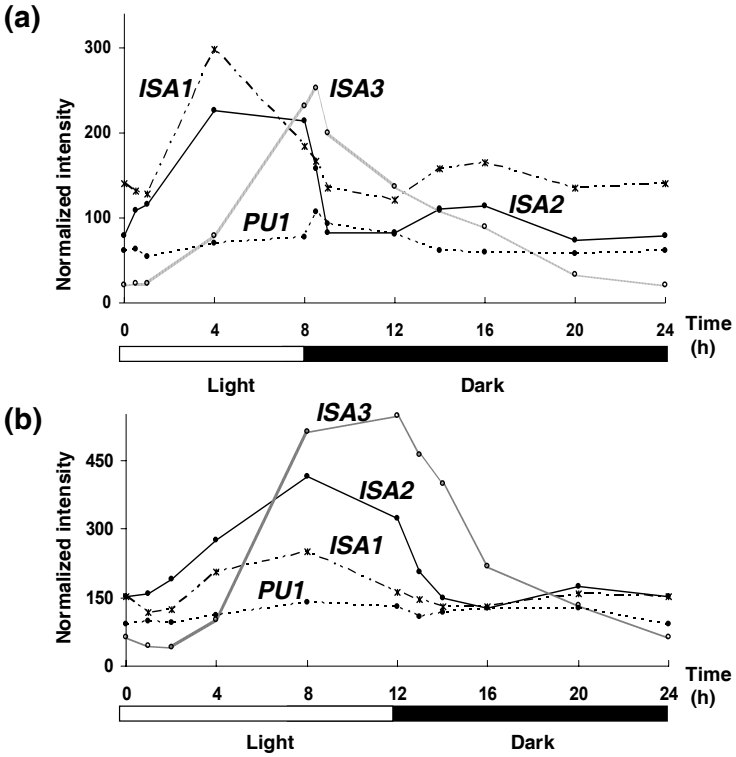


Figure 5. DBE have similar RNA accumulation patterns (WT) over different diurnal cycles.
(a) Under SD conditions.
(b) Under a 12 h light/12 h dark cycle.
The average normalized intensity for each chip is 100.

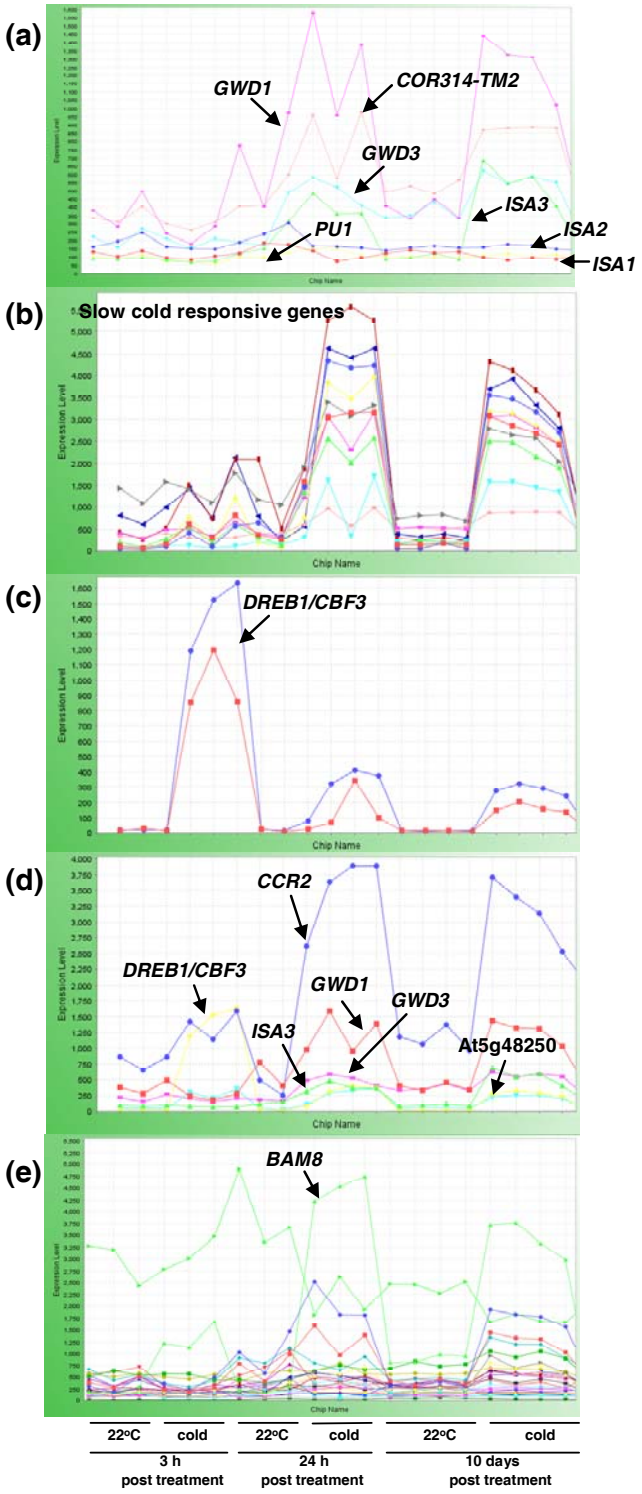

(a)
GWD1
COR314-TM2
GWD3
ISA3
PU1
ISA2
ISA1
(b)
Slow cold responsive genes
Expression Level
Chip Name
(c)
DREB1/CBF3
(d)
CCR2
DREB1/CBF3
GWD1
GWD3
ISA3
At5g48250
(e)
BAM8
22°C
cold
22°C
cold
22°C
cold
3 h
post treatment
24 h
post treatment
10 days
post treatment

Figure 6.

Figure 6. (continued). RNA accumulation profiles of starch metabolism genes and cold responsive genes in MetaOmGraph.
RNA accumulation of DBE plus *GWD1* and *GWD3* and a slow cold responsive gene (a), slow cold responsive genes (b), two *DREB1* genes (c), *ISA3*, *GWD1*, *GWD3* and the transcription factors that have RNA-accumulation correlation coefficient 0.5 or above with them (d), starch metabolic genes that have RNA-accumulation correlation coefficient no less than 0.5 with slow cold responsive genes (e). They are plotted from three cold experiments in the order of 3 h, 24 h and 10 days post cold treatment. There are 6, 6 and 8 chips in the three experiments. The first and fourth,7^{th} and 10^{th}, 13^{th} and 17^{th} chips are from WT; others are from cold sensitive mutants. Within each experiment, chips under the same condition are grouped together with control first; the genotypes are in the same order under each condition. The average expression level for each chip is 100.

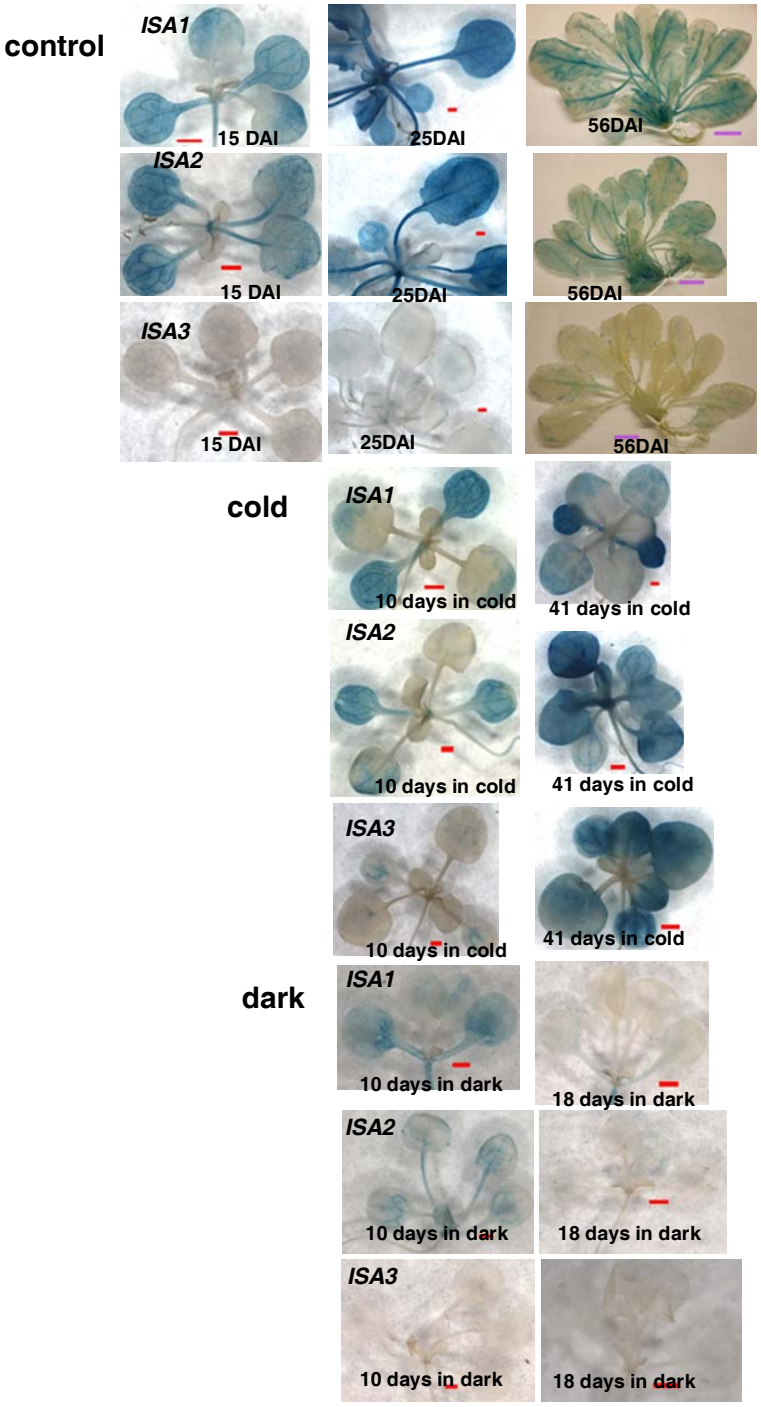


Figure 7. DBE expression is altered under changed conditions. Cold and dark treatment. Red bar=1mm, purple bar=1cm.

Table 1. Genes with related functions with similar co-expressed transcript profiles have high correlation coefficients from MetaOmGraph.

	ISA1	*ISA2*	*ISA3*	*PU1*	*GWD1*	*GWD3*
ISA1	1	0.74	0.42	0.59	0.52	0.52
ISA2	0.74	1	0.58	0.73	0.67	0.62
ISA3	0.42	0.58	1	0.53	0.8	0.76
PU1	0.59	0.73	0.53	1	0.7	0.72
GWD1	0.52	0.67	0.8	0.7	1	0.85
GWD3	0.52	0.62	0.76	0.72	0.85	1

Matrix of RNA-accumulation Pearson correlation of starch metabolic genes (DBE and GWD).

Table 2. Promoter analysis for DBE and GWD genes.

Motif	Function	In gene	Sequence	Prom's bound in genome	P-value
CCA1 binding site motif	Responses to light	1 2 3 P G3	AA(A/C)AATCT	27%	0.03
Ibox promoter motif	Light responsive element	1 2 P G3	GATAAG	40%	0.093
ARF binding site motif	Auxin response factors,light	1 2 3 G1	TGTCTC	40%	0.076
Evening Element	Circadian control	3 P G1 G3	AAAATATCT	6%	<0.001
CACGTGMOTIF /G-box	Light	1 3 G1	CACGTG	15%	0.035
G-box	Light	1 3	GCCACGTG(C/A)A	N/A	N/A
TCT-motif	Light	1	TCTTAC	N/A	N/A
CARGCW8GAT	Light and circadian cycle motif	1 G3	CATTTATTAG	59%	0.923
ATHB2 binding site motif	Inhibits light response,	1	TAATAATTA	10%	0.467
ATCT-motif	Light	3	TTCGATT	N/A	N/A
BoxII promoter motif	Light	P	GGTTAA	42%	0.933
GATA-motif	Light	P	GATAGGA	N/A	N/A
GT1-motif	Light	P	GGTTAA	N/A	N/A
GBF1/2/3 BS in ADH1	Binds to the G-box, light	G1	CCACGTGG	1%	0.077
GBOXLERBCS	Light	G1	CCACGTGGC	2%	0.132
Box II	Light and circadian	G3	GGCCCA	N/A	N/A
ATCTA element	Photosynthesis-related	G3	TAGATATATAGAT	N/A	N/A
DREB1A/CBF3	Cold-stress response	3	ACCGACCT	7%	0.304
CARrG promoter motif	Low temperature	G1	CCAAATAAGG	7%	0.315
MYB4 binding site motif	Cold stress	2 3 P	AAC(T/A)AAC	75%	0.871
LTRE promoter motif	Low temperature, cold	3	ACCGACA	5%	0.229
DRE core motif	Dehydration- and low-temperature stress-inducible	3	(G/A)CCGAC	23%	0.706
MYB1AT	Stress response	1 3 P G1 G3	AAACCA	85%	0.505
AtMYC2 BS in RD22	Transcriptional activators in abscisic acid signaling	1 2 G1 G3	CACATG	35%	0.066
ABFs binding site motif	ABA	1 3 G1	CACGTGGC	3%	<0.001
ABRE binding site motif	ABA	1 3 G1	CACGTGGC	4%	0.001
LEAFYATAG	ABA	1 3	CCAATGT	10%	0.082
ACGTABREMOTIFA2OSEM	ABA	1 3 G1	ACGTGGC	14%	0.032
ABRE-like binding site motif	ABA regulation	1 3 G1	CACGTGGC	20%	0.078
ABRE	ABA	P	TACGTG	N/A	N/A
W-box promoter motif	WRKY class binding site, senescence and plant defense	1 2 3 G1	TTGACT	67%	0.503
MYCATERD1	Water stress	1 2 G1 G3	CATGTG	35%	0.066
MYB2AT	Water stress	1 G1	TAACTG	29%	0.432
MBS	Water stress	1 3	(T/C)AACTG	N/A	N/A
GAREAT	GA	1 P G1 G3	TAACAAG	55%	0.3
CAAT-box	Cis-acting element in promoter and enhancer	1 2 3 P G1 G3	GGCAAT	N/A	N/A
SV40 core promoter motif	Activate transcription	P G1	GTGG(A/T)T(A/T)G	20%	0.23
LS5 /TGA2	Negative regulatory	1 2	T(A/C)TACGTCAC	N/A	N/A
MYB1LEPR	Regulation of cell cycle	2 3	GTTAGTT	17%	0.211
T-box promoter motif	Development	3 P G1 G3	ACTTTG	55%	0.271
TELO-box promoter motif	Required for the activation of expression in root primordia	2 P	AAACCCTAA	10%	0.084
L1-box promoter motif	Drives epidermal expression regulates epidermal patterning	1	TAAATG(T/C)A	14%	0.571
RY-repeat promoter motif	Seed	1	CATGCAT	3%	0.158
CAT-box	Meristem	1	GCCACT	N/A	N/A
CCGTCC-box	Meristem	G1	CCGTCC	N/A	N/A
HD-Zip1	Palisade mesophyll cell	2	CAAT(A/T)ATTG	N/A	N/A
HD-Zip2	Leaf morphology	2	CAAT(G/C)ATTG	N/A	N/A
TATA-box Motif	Promoter element around -30 of transcription start	1 2 3 P G1 G3	TATAAA	82%	0.185

In the category of "In gene", 1-3 stand for *ISA1-3*; P stands for *PU1*; G1 stands for *GWD1*; G3 stands for *GWD3*. Promoter motifs of interest are marked with color: light responsive promoter motif, cold responsive promoter motif. P-value (<=0.1) from promoter motif of interest is also marked. The information mostly comes from Athena, and also Plant Care.

Table 3. Starch metabolism genes with RNA accumulation profiles tightly correlated to DBEs and GWDs. Their transcript profiles are highly correlated to that of temperature responsive genes.

Gene	*ISA1*	*ISA2*	*PU1*	*PHS1*	*DPE1*	*PHS2*	*DPE2*	*MEX1*	*ISA3*	*GWD1*	*GWD3*	*SEX4*	*AAM3*	*BE3*	*BAM2*	Locus ID
ISA1	1.00	0.74	0.59	0.66	0.72	0.58	0.62	0.68	0.42	0.52	0.52	0.61	0.54	0.34	0.48	AT2G39930
ISA2	0.74	1.00	0.73	0.78	0.64	0.73	0.80	0.75	0.58	0.67	0.62	0.72	0.63	0.46	0.68	AT1G03310
PU1	0.59	0.73	1.00	0.66	0.59	0.71	0.65	0.79	0.53	0.70	0.72	0.66	0.72	0.48	0.66	AT5G04360
PHS1	0.66	0.78	0.66	1.00	0.83	0.72	0.74	0.68	0.53	0.67	0.64	0.66	0.79	0.46	0.46	AT3G29320
DPE1	0.72	0.64	0.59	0.83	1.00	0.71	0.66	0.68	0.56	0.67	0.68	0.71	0.79	0.50	0.40	AT5G64860
PHS2	0.58	0.73	0.71	0.72	0.71	1.00	0.89	0.70	0.81	0.92	0.83	0.89	0.78	0.67	0.68	AT3G46970
DPE2	0.62	0.80	0.65	0.74	0.66	0.89	1.00	0.67	0.79	0.85	0.74	0.85	0.72	0.62	0.66	AT2G40840
MEX1	0.68	0.75	0.79	0.68	0.68	0.70	0.67	1.00	0.52	0.69	0.74	0.75	0.71	0.51	0.69	AT5G17520
ISA3	0.42	0.58	0.53	0.53	0.56	0.81	0.79	0.52	1.00	0.80	0.76	0.82	0.53	0.74	0.56	AT4G09020
GWD1	0.52	0.67	0.70	0.67	0.67	0.92	0.85	0.69	0.80	1.00	0.85	0.82	0.79	0.68	0.69	AT1G10760
GWD3	0.52	0.62	0.72	0.64	0.68	0.83	0.74	0.74	0.76	0.85	1.00	0.81	0.72	0.63	0.72	AT5G26570
SEX4	0.61	0.72	0.66	0.66	0.71	0.89	0.85	0.75	0.82	0.82	0.81	1.00	0.71	0.67	0.68	AT3G52180
AAM3	0.54	0.63	0.72	0.79	0.79	0.78	0.72	0.71	0.53	0.79	0.72	0.71	1.00	0.47	0.56	AT1G69830
BE3	0.34	0.46	0.48	0.46	0.50	0.67	0.62	0.51	0.74	0.68	0.63	0.67	0.47	1.00	0.49	AT2G36390
BAM2	0.48	0.68	0.66	0.46	0.40	0.68	0.66	0.69	0.56	0.69	0.72	0.68	0.56	0.49	1.00	AT2G32290
small Hsp	0.57	0.67	0.70	0.61	0.59	0.82	0.76	0.75	0.61	0.82	0.74	0.77	0.76	0.48	0.73	AT1G06460
Hsp	0.35	0.44	0.52	0.38	0.47	0.72	0.60	0.59	0.64	0.74	0.74	0.71	0.60	0.50	0.70	AT5G23240
LTI6A	0.30	0.43	0.53	0.54	0.58	0.78	0.63	0.54	0.63	0.74	0.61	0.73	0.73	0.53	0.44	AT4G30650
COR15B	0.27	0.44	0.55	0.46	0.52	0.76	0.64	0.53	0.62	0.72	0.58	0.71	0.68	0.56	0.41	AT2G42530
KIN1	0.28	0.45	0.58	0.36	0.42	0.67	0.55	0.59	0.64	0.67	0.56	0.66	0.55	0.53	0.57	AT5G15960
COR15A	0.27	0.40	0.51	0.33	0.41	0.67	0.55	0.54	0.58	0.66	0.52	0.67	0.59	0.51	0.44	AT2G42540
COR314-TM2	0.56	0.62	0.72	0.51	0.53	0.72	0.64	0.81	0.53	0.70	0.64	0.73	0.66	0.45	0.65	AT1G29390
COR414-TM1	0.35	0.47	0.57	0.36	0.41	0.72	0.62	0.64	0.61	0.71	0.59	0.71	0.56	0.51	0.58	AT1G29395

Matrix of correlation of genes with Pearson correlation no less than 0.5 to *ISA3*, *GWD1* and *GWD3*. *ISA1* is included also. Genes involved in starch metabolism are marked green; genes related to heat shock are marked pink; genes related to cold response are marked blue. Correlation coefficients >= 0.85 are marked maroon, >= 0.75 are marked orange, >=0.65 are marked yellow, >=0.55 are marked light yellow.

CHAPTER 3. UNRAVELING THE CONNECTION BETWEEN STARCH GRANULE METABOLISM AND ACETYL-COA UTILIZATION

A paper to be submitted to Plant Physiology

Ling Li[1], Carol M. Foster[1], Wieslawa I. Mentzen, Basil J. Nikolau, Dan Nettleton, Martha G. James, Alan M. Myers, and Eve SyrkinWurtele

[1]These authors contributed equally to this paper.

ABSTRACT

Starch accumulates and degrades in leaf chloroplasts during the light and dark phases of the diurnal cycle, providing a stored source of energy for plant growth and metabolism. This fluctuation in starch content is also integrated with the overall physiological state of the plant. We explore the starch metabolic network in the context of other metabolic processes by investigating mutants with single-gene changes that perturb starch metabolism. The central metabolite, acetyl-CoA, is required for many biochemical pathways critical to plant growth and development. ATP citrate lyase (ACL) produces acetyl-CoA from citrate in the cytosol. Transgenic Arabidopsis with decreased levels of ACL activity contain abnormally enlarged starch granules in leaves under long day (LD) photoperiod. To study the underlying molecular events involved in utilization of acetyl-CoA during starch metabolism in wild-type (WT) and antisense-*ACLA* plants, leaves of seedlings grown under short day (SD) photoperiod were analyzed by microarray, and starch was characterized. Zymogram analyses of starch metabolizing enzymes were used to compare enzyme activity to transcript accumulation profiles. The antisense-*ACLA* mutant accumulates more starch than WT under SD conditions. There are differences between WT and antisense-*ACLA* in the context of differential expression of genes implicated in starch, acetyl-CoA metabolism and other metabolic pathways, and a tight correlation between accumulation of RNAs of starch synthesis and degradation.

INTRODUCTION

In plants, storage of carbohydrate in starch granules provides a mechanism by which photosynthate can be remobilized to sustain respiration and sucrose metabolism for growth and development (Gibon et al., 2004). The complex regulatory mechanisms that direct the net accumulation of starch and sugars, and the integration of this carbohydrate network into central metabolism are key to the flow of carbon *in planta*. In leaves, diurnal synthesis and use of starch and sugars is thought to mitigate changes in the net pool of carbon **(**Fox and Geiger, 1984**)**, however, many questions remain about the mechanisms of starch granule production and degradation, or its integration with central metabolism, in this organ. Ultimately, clarifying the process of granule biosynthesis and unraveling interactions between starch metabolism and other metabolic pathways may broaden the functionality of starch for human use.

Starch granules accumulate and degrade in leaf chloroplasts during the light and dark, respectively. Only a small amount of starch is retained at the end of a dark period in unstressed plants, regardless of light conditions (Geiger and Servaites, 1994; Matt et al., 1998; Zeeman et al., 1998). More starch accumulates during a short day (SD) photoperiod than under long day (LD) conditions (Matt et al., 1998), providing more carbon resources during a longer night (Gibon et al., 2004). These data indicate formation of granules is highly regulated throughout the diurnal cycle.

Starch granules are highly-organized, densely-packed, semicrystalline structures composed of two different glucan chains, amylose and amylopectin (Myers et al., 2000). Amylose is a linear, unbranched polymer of α-(1→4)-linked glucose units, where-as, amylopectin contains α-(1→4) and α-(1→6) bonds resulting in numerous branched chains. Although both polymers are found in storage and transient granules, the primary component is amylopectin. Its ordered structure and clustered arrangement within the granule play crucial roles in starch function; not only starch level, but also granule composition can be altered by changes in diurnal conditions (Zeeman et al., 2002). In Arabidopsis, the proportion of amylose in wild-type (WT) starch granules increased from 4% to 25% after the light period was extended from a 12-h photoperiod to continuous light (CL).

Accumulation of starch is mediated by the metabolic status of the plant and is extremely sensitive to external and endogenous stimuli. In leaves, levels of glucose and sucrose (Koch, 1996), light (Neff et al., 2000), and the circadian clock (Harmer et al., 2000) influence regulatory mechanisms that control gene expression implicated in balancing production and export of carbohydrate with consumption. Turnover of carbohydrates in different tissues and organs in response to environmental fluctuations and nutrient availability is integrated by this regulatory network, thus, maintaining carbon balance throughout the plant (Baier et al., 2004; Coruzzi and Bush, 2001).

One approach to understanding starch metabolism and its regulation in the context of other metabolic processes is to investigate mutants with single-gene changes that perturb starch metabolism. One such mutant involves the ATP citrate lyase (ACL) gene (Fatland et al., 2005). The central metabolite, acetyl-CoA, is required for many biochemical pathways crucial to plant growth and development. ACL produces acetyl-CoA from citrate in the cytosol (Fatland et al., 2002). Transgenic antisense-*ACLA* Arabidopsis with decreased levels of ACL activity exhibit a distinct phenotype and contain abnormally enlarged starch granules in leaves under LD conditions (Fatland et al., 2005) indicating an imbalance between starch biosynthesis and mobilization.

To elucidate underlying molecular events involved in starch granule metabolism and its connection to acetyl-CoA utilization, we analyzed starch accumulation, enzyme activities, and global gene expression patterns in WT and antisense-*ACLA* plants grown under a SD photoperiod. Our objectives were to examine activity of starch metabolism genes and to determine how decreased cytosolic ACL activity affects pathways involved in carbon metabolism and may lead to subsequent changes in metabolic flux throughout the plant.

RESULTS

Phenotype of Antisense-*ACLA* Plants Grown under SD Conditions

The antisense-*ACLA* plants grown under LD have an extreme miniature phenotype, with very tiny dark leaves and delayed senescence, as well as high starch (Fatland et al., 2005). To identify genes related to starch accumulation, independent of genes associated with other pleotropic effects of the ACL mutation, we wished to identify conditions in which the

high-starch phenotype of the mutants was maintained, but other aspects of the phenotype were minimized. Initial experiments varying day length indicated that the light conditions show a strong effect on the penetrance of the phenotype of antisense-*ACLA* plants. To identify light conditions under which the overall phenotype of antisense-*ACLA* and WT plants is similar, yet the level of starch accumulation is altered, plants were grown under various conditions including SD. At 42 DAP, the phenotype was accessed, and starch levels were quantified. Under SD, the overall phenotype of the antisense-*ACLA* and WT plants is quite similar, although leaves and rosette diameter are smaller (Figure 1). Bolting in both antisense-*ACLA* plants and WT plants is delayed similarly in response to the decreased day-length. However, starch accumulation is increased in the SD-grown antisense-*ACLA* plants. At the end of the light period, antisense-*ACLA* plants (4.0 ± 0.2 mg starch g^{-1} fresh weight) had more than double the starch content of wild-type plants (1.8 ± 0.1 mg starch g^{-1} fresh weight). The difference between mean starch content was significant (t test; $P < 0.001$).

Leaves were examined over a diurnal cycle. Plants were harvested at 6 weeks (i.e., before initiation of bolting) to maximize physiological similarity across plants. Leaves 5 to 8 (staged according to Bowmann, 1994) were harvested; these leaves are mature and photosynthetically active (Stessman et al., 2002). Leaves from 16 plants were harvested and pooled for each sample. The experiment was run in duplicate. Plants were harvested at 12 time points throughout the day-night cycle, at 30 min and 1 h after onset of light and onset of dark to enable identification of genes that are expressed quickly upon changes in the light exposure; additional harvests were at 4-h intervals.

Expression of Circadian Rhythm Genes under SD Conditions

To characterize the responses of known genes with circadian rhythms in the short day diurnal cycle, we monitored the RNA accumulation of light harvest complex (Lhcb3) and chalcone synthase (CHS), two light-regulated genes with known diurnal or circadian rhythms under LD (12 h of light followed by 12 h of dark). Total RNA was isolated from the harvested leaf samples from the 12 time points (0, 0.5, 1, 4, and 6 h in the light, 0, 0.5, 1, 4, 6, 8, and 12 h in the dark). Lhc genes peak around midday and anticipate light at the end of night (Harmer et al., 2000). Northern blot analyses of WT plants for *Lhcb3* and *CHS*

transcript showed *Lhcb3* and *CHS* both display diurnal rhythms under SD and the circadian behavior of *Lhcb3* and *CHS* in the antisense-*ACLA* mutant was similar to WT (Figure 2).

Fluctuation in Transcriptome of WT under SD Diurnal Cycle

Light plays an important role in regulating plant growth and development. Various processes occur over the diurnal cycle. For example, starch is accumulated in the light and degraded in the dark. Smith et al. (2004) reported the diurnal changes in the transcriptome of starch metabolic genes under a 12 h light/12 h dark condition. Northern blot analysis of six genes showed that expression patterns were well characterized by 11 of the 12 time points because the 6 h in light time point consistently fell near a line segment connecting the adjacent points. To identify the expression of the genes that might have different temporal pattern over the SD diurnal cycle (especially the starch metabolic genes), the transcript abundance was monitored at 11 time points (time point 6 h in the light was not selected based on the northern blots analysis) with two biological replicates. As described in "Materials and Methods", the RNA extracted from the leaves was analyzed using Affymetrix Arabidopsis ATH1 arrays (with 22,810 Arabidopsis probe sets).

Gene expression was compared over the 11 time points in WT. Four hundred and seventeen genes were significantly differently expressed at the level of 0.00001 based on the p-values. Two hundred and seventy genes were significant when the false discovery rate (FDR) was controlled at the level of 0.00031 using the method of Benjamini and Hochberg (1995). According to their different accumulation patterns, these genes were grouped into 7 clusters using K-medoids clustering approach (see Materials and Methods) (Figure 3 and Appendix E, Table S1). Genes within each cluster have a similar expression pattern. Except Cluster 4, all other clusters have quite a few genes that are annotated as "expressed protein" and "functions unknown". Genes in Cluster 1 have peaks at the onset of light and lowest expression at the beginning of dark. Genes in Cluster 2 have peaks at the beginning of light (half an hour to 1 h in the light), and lowest expression at the beginning of dark (1 h in the dark). Genes in Cluster 3 also peak at the beginning of light, but they have a delayed lowest expression time (at 4 h in the dark). Genes in Cluster 4 have peaks in the light at about 4 h and lowest expression at 4 h to 8 h in the dark. Genes in Cluster 5 have peaks at 0.5 h in the

dark; their expression decreases rapidly in the dark and increases rapidly in the light. Compared to the genes in Cluster 5, genes in Cluster 6 increase their expressions more slowly. They have peaks at 1 h in the dark and decrease their expression more slowly. Genes in Cluster 7 have high expression in the dark and low expression in the light. Clusters with different diurnal patterns of expression may represent different biological processes corresponding to the demands of plant growth and development. In the clusters with peaks in the light, the proportions of the genes with sub-cellular localization in the chloroplast are higher than those in the dark, e.g., 23.1% in Cluster 4 (in which gene expression is high in the light) and 11.1% in Cluster 7 (in which gene expression is high in the dark); while in the clusters with peaks in the dark, the proportions of the genes with sub-cellular localization in the mitochondrion are higher than those in the light, e.g., 7.4% in Cluster 7 but no gene in Cluster 4 is predicted to be localized in mitochondrion (Appendix E, Table S1). For the clusters with peaks in the light, Cluster 1 has several genes involved in regulation of transcription, and protein metabolism; Cluster 2 has more genes involved in regulation of transcription and some genes are involved in transport; Cluster 3 has decreased number of genes involved in regulation of transcription, but more genes are involved in protein metabolism and still quite a few genes are involved in transport; Cluster 4 has a wide range of peak time in the light and it contains fewer genes involved in regulation of transcription, but several genes are involved in secondary metabolism. For the clusters with peaks in the dark, Cluster 5 has genes involved in starch and sucrose degradation, several cold responsive genes, and some genes involved in hormone metabolism; Cluster 6 contains more genes involved in regulation of transcription and cold response, and genes involved in stress response or hormone metabolism; Cluster 7 contains several genes involved in regulation of transcription or cell cycle or development.

Of particular interest, the 270 genes designated to have the most significant change (FDR<0.00031) in expression across the diurnal cycle include multiple genes that are known to respond to cold (Cluster 5 and 6, with peaks at the beginning of dark; Appendix E, Table S1). Expression of genes of the DREB1 and DREB2 families (except DREB1C), which can activate cold-regulated genes or induce drought/osmotic stress genes (Liu et al., 1998), also have peaks at the beginning of dark (data not shown).

Expression of Genes of Starch Biosynthesis and Degradation over the SD Diurnal Cycle

The mRNAs of genes encoding enzymes involved in starch metabolism accumulated in three overall patterns: peak in the light; peak in the dark; peak both in the light and in the dark (Figure 4). Figure 5 shows the accumulation profiles of starch metabolism RNAs, grouped by their biochemical activity.

Starch synthases (SS) transfer glucosyl units from ADP-glucose to growing chains via new α-(1→4) linkages (Myers et al., 2000). The patterns of accumulation of the starch synthase mRNAs are quite distinct from one another (Figure 5A**).** *AtGBSS* shows a circadian pattern during the day/night cycle (Tenorio et al., 2003). *AtGBSS* is the most highly expressed of any of the starch metabolic genes. The mRNA abundance of *AtGBSS* peaks early in the light phase, declines to near zero by the end of the light phase, and begins to accumulate again in the middle of the dark phase. *AtSS2* mRNA abundance follows a similar pattern. *AtSS4* mRNA accumulation is low overall, but peaks at the beginning of the dark phase. *AtSS1* mRNA accumulates in moderate abundance, peaking at the beginning of the dark phase, declining in the middle of the dark, and increasing toward the end of the dark and throughout the light phase. *AtSS6* mRNA, a gene which has no apparent plastid targeting sequence, and has an as yet undefined role in metabolism, accumulation is moderate across the light/dark cycle, with two peaks, one in the middle of the dark and one at the end of the dark. Lowest levels for *AtSS6* occur at the beginning of the dark and towards the end of the dark phase. *AtSS3* has low mRNA abundance throughout the entire light/dark cycle.

Starch branching enzymes (BE) hydrolyze α-(1→4) linkages and introduce α-(1→6) linkages into the polymers (Myers et al., 2000). *AtBE2* mRNA accumulation peaks in the middle of the light and, with the exception of a very small peak at the beginning of the dark period, diminishes over the dark phase of the light cycle (Figure 5B**)**. Unlike the behavior of its mRNA accumulation under LD with a peak at 4 h before dark (Smith et al., 2004), under SD, *AtBE3* expression has a peak at 0.5 h in the dark. Very low accumulation of *AtBE1* is detected in the leaves.

Starch debranching enzymes (DBE) hydrolyze α-(1→6) linkages (Myers et al., 2000)**.** *AtISA1*, *AtISA2*, and *AtISA3* encode isoamylase-type DBEs, and *AtPU1* encodes pullulanase-type DBE. Their function in starch metabolism is still not clear. The profiles of *AtISA1* and

AtISA2 are similar (Figure 5C**)**. Both peak once in the middle of the light and once in the middle of the dark. The mRNA abundance peak in the light phase is nearly double that of the peak in the dark phase. *AtISA3* differs from *AtISA1* and *AtISA2* in that it has one, not two, mRNA peaks, which is at the beginning of the dark phase (0.5 h in the dark), unlike the behavior of its mRNA accumulation under LD with a peak at 4 h before the dark (Smith et al., 2004). The lowest levels of *AtISA3* mRNA accumulation occur during the end of the dark and beginning of the light phase. *AtPU1* is the DBE isoform with the lowest mRNA abundance. *AtPU1* peaks slightly at the beginning of the dark phase. It decreases over the first half of the dark phase, then begins to accumulate again during the end of the dark phase and during the light phase.

Starch disproportionating enzymes (DPE) transfer residues between linear chains by the cleavage and reformation of α-(1→4) linkages, and less amylopectin is accumulated when DPE is missing (Myers et al., 2000). *AtDPE2* mRNA accumulation peaks at the onset of the dark phase (Figure 5D**)**, unlike the behavior of its mRNA accumulation under LD with a peak at 4 h before the dark (Smith et al., 2004). Following this peak, mRNA accumulation drops quickly, and then diminishes slowly over the rest of the dark phase. *AtDPE1* mRNA also peaks at the beginning of the dark phase, but at a much lower level than *AtDPE2* mRNA.

Beta-amylases (BAM) are α-(1→4)-specific hydrolases, working from the nonreducing end; the degradative action stops at each branch linkage. Although the primary function for BAMs is thought to be starch breakdown (Smith et al., 2005), the elevated transcript levels during both light and dark periods suggest there might be alternative functions for AtBAM isoforms in leaves (Figure 5E**)**. *AtBAM8* has the most mRNA accumulation of any gene in the BAM study set. It has three peaks of accumulation, in the early light phase, the beginning of the dark phase, and the later hours of the dark phase. Unlike the behavior of its mRNA accumulation under LD with a peak at the onset of dark (Smith et al., 2004), under SD, *AtBAM8* expression has a peak at the beginning of dark (0.5 h in the dark). The *AtBAM3* mRNA accumulation pattern is similar. *AtBAM7* mRNA accumulation gently peaks in the early light phase and again in the middle of the dark phase. *AtBAM9* mRNA accumulation peaks at the end of the dark and beginning of the light phase,

with lowest levels at the end of the light phase. Five *AtBAM*s (1, 2, 4-6) accumulate only to very low levels through the diurnal cycle.

Alpha-amylases (AAM) are also α-(1→4)-specific hydrolases; unlike the beta-amylases, they cleave within linear chains. Their putative role during starch metabolism is to attack the starch granule surface and release soluble glucans for further degradation (Smith et al., 2005). *AtAAM3* mRNA accumulation is the highest of the AAM family (Figure 5F**)**. Unlike the behavior of its mRNA accumulation under LD with a peak more toward the light (Smith et al., 2004), .under SD, *AtAAM3* mRNA accumulation peaks more toward the dark. *AtAAM1* mRNA levels are lowest when *AtAAM3* levels are highest. *AtAAM2* has very low levels of mRNA accumulation over the entire light cycle.

Starch phosphorylases (PHS) insert phosphoryl groups from inorganic pyrophosphate into α-(1→4) bonds of starch polymers, releasing glucose-1-phosphate. PHSs function in starch degradation (Smith et al., 2005). *AtPHS2* mRNA accumulation increases throughout the light phase and reaches relatively high levels at the onset of the dark phase, when starch granule disassembly begins, as compared to other genes in the study set (Figure 5G**)**. *AtPHS2* mRNA levels diminish slowly over the dark phase, reaching their lowest at the end of the dark. *AtPHS1* mRNA accumulation also peaks at the end of the light/beginning of the dark phase, but is unique in that it dips slightly just as it peaks. Unlike the behavior of the two *PHS*s mRNA accumulation under LD with a peak more toward the light (Smith et al., 2004), .under SD, their mRNA accumulation peaks more toward the dark.

Of the other genes related to starch metabolism, glucose-6-phosphate isomerase (*PGI1*) peaks in the light (Figure 5H**)**. Glucan water dikinase (*GWD1*) is highly expressed early in dark. Phosphoglucomutase (*PGM1*) peaks early in the dark. ADP-glucose pyrophosphorylase (*ADPG-PP*), an allosteric, heterotetrameric enzyme composed of two large and two small subunits, catalyzes the synthesis of the starch precursor molecule, ADP-glucose, and also plays an important regulatory role in controlling starch biosynthesis (Hendriks et al., 2003; Baroja-Fernández et al., 2004; Cross et al., 2004). A small subunit encoding gene *APS1* and a large subunit encoding gene *APL1* are highly expressed with *APL1* peaking in the light and *APS1* peaking in the dark. Other large subunit encoding gene

APL2, *3*, and *4* are expressed very lowly. This result is consistent with the observation that the APS1 (catalytic)/APL1 (allosteric) complex has highest activity (Crevillén et al., 2003).

RNAs whose accumulation peaks soon after the shift to light include *AtGBSS*, and, somewhat counter intuitively, the starch catabolic genes *AtBAM8*, *AtBAM3*, *AtBAM9* and *AtAAM1*. RNAs whose accumulation peaks in the middle of light phase include *AtBE2*, *AtISA1* and *AtISA2*, and *AtPGI1*. Genes whose transcript accumulation peaks during the light period may play a role in starch synthesis. In contrast, *AtBE3*, *AtISA3*, *AtPU1*, *AtDPE2*, *AtBAM8*, *AtBAM3*, *AtAAM3*, *AtPHS1*, *AtPHS2*, *AtAPS1*, *AtPGM1*, and *AtGWD1* peak soon after the shift to dark. *AtISA1*, *AtISA2*, *AtBAM8*, *AtBAM9*, and *AtAPS1* peak 8 h or more after the shift to dark. Genes whose accumulation peaks in the dark may play a role in the starch degradation. Of the genes whose RNA accumulation peaks in the dark, *AtBAM3*, *AtBAM8*, and *AtBAM9* have peaks both in the dark and in the light. They might function both in the starch biosynthesis and degradation.

Enzyme Activity Does Not Reflect Transcript Accumulation Profiles

Smith et al. (2004) and Lu et al. (2005) have reported the levels of three proteins involved in starch degradation (DPE2 (At2g40840), AMM3 (At1g69830) and GWD1 (At1g10760)) are constant in spite of the fluctuation of the corresponding transcript. To compare enzyme activity to transcript accumulation profiles for the DBEs, BEs, AAMs, and the soluble SSs, and to study the effects of the dominant antisense ACL allele on the activity of these enzymes, we used native PAGE to separate proteins from extracts of photosynthetically active rosette leaves. After electrophoresis, DBE, BE, BAM, and AAM activities were assayed by transferring proteins to polyacrylamide gels containing solubilized potato starch, incubating the gels and staining with I_2/KI (Dinges et al., 2001). Relative migration distances together with changes in the dark blue color of I_2/KI-stained starch in these gels allowed for detection and provisional identification of the enzymes (Kakefuda and Duke, 1984).

Six major bands were visible that represented two putative debranching enzymes (AtISA2 and AtPU1) and three putative branching enzymes (AtBE1, AtBE2, and AtBE3) (Figure 6A). In addition, numerous clear bands were present that may denote amylase

activities. In all cases, activity of each protein appears to be relatively uniform across all time points, thus, enzyme activity represented by each band does not reflect the fluctuations observed in transcript accumulation in response to the diurnal cycle (Figure 6A).

A single amylase band, discerned at time points 4 h in the light and 0 h in the dark, peaking at 0.5 h after onset of dark as a colorless band at the bottom of the gel, distinguishes antisense-*ACLA* from WT plants (Figure 6A). This band is undetectable in WT.

To detect soluble starch synthase activities, native PAGE was conducted in a gel containing glycogen (Cao et al., 1999). Gels were incubated with ADP-glucose to provide substrate, and stained with I_2/KI. Dark blue bands in the light brown-colored gel are due to starch, indicating the presence of starch synthase activity. Multiple colorless bands, presumably representing glycogen-hydrolyzing enzymes, are also present. Three starch synthase isoforms were resolved from both WT and antisense-*ACLA* leaves (Figure 6B). The relative migration distances of these isozymes are comparable to those of AtSS3 and AtSS1 (personal communication-French research group), the most active of the soluble SSs in leaves. Activity patterns of starch synthases in WT and antisense-*ACLA* were similar and not noticeably changed by the diurnal cycle (Figure 6B).

At the end of the light period (4 h in the light), two colorless bands located between the largest colorless band and the large AtSS1 band were visible in antisense-*ACLA*, but not in WT; these bands persisted in the antisense-*ACLA* plants throughout the dark period (0 h, 0.5 h, 1 h, and 4 h in the dark) (Figure 6B). Thus, starch modifying activity is altered in the antisense-*ACLA* mutation.

Co-expression of Starch Metabolic Genes across Multiple, Diverse Microarray Experiments

To better understand the transcriptional regulation of starch metabolism, our goal is twofold. First, to identify starch metabolism genes which consistently cluster together across diverse experiments to give us clues as to the mechanisms of regulation of the starch network. Second, to identify known and unknown genes in those clusters to help elucidate the mechanisms involved in starch granule production and degradation. To achieve this goal, profiles for starch metabolism genes from a co-normalized set of public microarray data

(Mentzen, 2006) were compared across the nearly 1000 chips representing 70 Arabidopsis Affymetrix ATH1 microarray experiments available in the Nottingham Arabidopsis Stock Centre microarray (NASCArray) database (http://affymetrix.arabidopsis.info/narrays/experimentbrowse.pl, Craigon et al., 2004). These experiments examine various mutants, organ types, development stages, and responses to biotic and abiotic stresses, such as pathogens and light conditions. To improve our ability to custom analyze data from large numbers of microarray chips, and bring in metadata information, we have developed a new open source software component of the MetNet platform, MetaOmGraph (MOG, http://www.metnetdb.org/MetNet_MetaOmGraph.htm).

Because *ISA1* (At2g39930) plays an important role in starch biosynthesis, possibly as a member of a protein complex with *ISA2* (At1g03310) (Hussain et al., 2003), and because under SD conditions, RNA accumulation for both *ISA1* and *ISA2* peaked during the light phase of the diurnal cycle, we chose to examine the co-expression group associated with this gene. Our hypothesis was that levels of *ISA1* and *ISA2* RNAs would be extensively correlated throughout development and across environmental perturbations. Our second hypothesis was that this type of co-expression analysis would reveal additional genes of starch synthesis that are highly correlated with *ISA1*. Using Pearson correlation as a similarity measure, we searched for genes with expression patterns that are similar to *ISA1* across the experiments in the NASCArray database. In these experiments, we evaluated co-expression between *ISA1* and the 22,746 genes on the ATH1 chip.

Eight genes were found to be highly correlated (0.66 to 0.74) to *ISA1* (Table 1). Interestingly, three of these genes, *DPE1* (At5g64860), *PHS1* (At3g29320), and ROOT CAP 1 (*RCP1*, At5g17520), participate in starch degradation (Smith et al., 2003, Niittylä et al., 2004). *RCP1*, previously known as *MEX1*, is a chloroplastic maltose transporter critical during starch degradation. *Pti1* (At1g26150) encodes a putative serine/threonine kinase with similarity to Pto kinase interactor 1 (Zhou et al., 1995). Pti1 participates in plant-pathogen interactions (Zhou et al., 1995; Bogdanove and Martin, 2000) and may be part of a signal transduction pathway. The strong RNA accumulation correlation between the putative serine/threonine kinase and *ISA1* suggests this kinase may play a role in phosphorylation during starch metabolism. Interestingly, gene expression of a CbbY-like protein

(At5g45170), predicted to be chloroplastic, closely correlates with *ISA1*. Although CbbY, a bacterial gene located in the fructose 1,6-bisphosphate aldolase operon of *Hydrogenophilus thermoluteolus* (Hayashi et al., 2000), has no assigned function, it is a member of a prokaryotic regulon sensitive to metabolites that signal the nutritional status of a cell (Bowien and Kusian, 2002). Other genes in the *ISA1* co-expression group are D-cysteine desulfhydrase (At1g48420), which catalyzes the breakdown of cysteine to produce pyruvate, ammonia, and sulfuric acid in mitochondrion (Riemenschneider et al., 2005), and two genes with unknown functions, mitochondrial At3g60810 and chloroplastic At4g10470.

Because *DPE2* mRNA accumulates during the dark phase of the SD photoperiod and plays an important role in starch breakdown (conversion maltose to sucrose) in the cytosol (Chia et al., 2004; Lu and Sharkey, 2004). We checked the correlation of *DPE2* with other genes in MOG. Expression profiles for nine genes (correlation coefficients of 0.72 to 0.84) were found to be similar to that of *DPE2* (Figure 7). Of these, six genes, *GWD1* (At1g10760), *GWD3* (At5g26570), *ISA2* (At1g03310), *ISA3* (At4g09020), *PHS1* (At3g29320), and *PHS2* (At3g46970), are also known to be involved in starch degradation. *DPE2* was most highly correlated with a protein tyrosine phosphatase/kinase interaction sequence protein (*PTPKIS1*/*SEX4*, At3g52180). PTPKIS1 is a novel tyrosine phosphatase that dephosphorylates proteins associated with the SnRK complex (Fordham-Skelton et al., 2002). Inactivation of the SnRK regulatory complex affects vital metabolic enzymes, such as sucrose synthase and sucrose phosphate synthase, and possibly alters global responses to carbon and stress signaling (Halford and Hardie, 1998). Our data are consistent with results reported by Schaffer et al. (2001) that used gene expression data from the Stanford Microarray Facility. These data also show that *PTPKIS1* and *PHS2* have similar expression profiles. The strong correlation of *PTPKIS1* expression profiles to those of the *DPE2* co-expression group of starch metabolism genes indicates a link between *PTPKIS1* and carbohydrate metabolism. Our hypothesis was confirmed by Niittylä et al. (2006). *SEX4* is involved in starch degradation.

To better identify the relationship among the RNAs in the starch metabolic network, we analyzed gene expression groups within this starch network across multiple conditions. Starch metabolic genes and putative regulators, *PTPKIS1* and Pti-like protein, were grouped

according to Pearson's correlation values and visualized in a color-coded matrix (Figure 8). Three distinct groups of co-expressed genes emerged. One group contains *DPE2*, *GWD1*, *GWD3*, *PHS2*, *SEX4*, *ISA3*, *BAM2* and *BE3*. The second group contains the starch synthesis genes (*ISA1* and *ISA2*), *DPE1*, *PHS1*, and the gene encoding Pti1-like protein.

In the third group of co-expressed genes, three of the eight genes [*ISA2*, *SS4*, and *APL1* (At5g19220)] play a role in starch synthesis. Two genes are linked to degradation, *PU1* and *AAM3* (chloroplastic, glycogen catabolism). The remaining three genes, triose phosphate/phosphate translocator gene, *TPT1* (At5g46110), phosphoglucomutase 1, *PGM1* (At5g51820), and a PGM-like4 (At1g70820), are critical to carbon flux.

Co-regulation of three pairs of genes, *AAM1* and PGM-like2 (At4g11570); *SS2* and AGL-like3 (At1g24320); and *BE1* and *BAM5*, was noted. Each of these three gene pairs has at least one member (*AAM1*, *SS2*, *BE1*, and *BAM5*) with peak transcript accumulation during the light period of the SD photoperiod.

Only five genes expected to play a role in starch metabolism (*BAM9*, *AAM2*, *BAM3*, *BAM6*, and *BAM1*) are not highly correlated to any other known starch metabolism genes or to each other. Five *AtBAM*s (1, 2, 4-6) accumulate only to very low levels throughout the diurnal cycle. The metabolic function of these genes is unknown.

Antisense-*ACLA* Shows Complex Altered Gene Transcripts

To investigate potential changes in gene expression that might contribute to the increase in starch in antisense lines, we profiled transcript levels in antisense-*ACLA* plants and compared them to the WT plants. The antisense-*ACLA* mutants had been planted together with their WT counter parts (split-plot experiment design), grown and harvested, RNA extracted for all 12 time points along with the WT plants with two replicates. Because of the diurnal fluctuations in gene expression, we selected different times in the diurnal cycle for this comparison. Specifically, we selected the following three time points: 1 h in the light, 0.5 h and 4 h in the dark based on the northern blots of *BE3*, *AAM2* and *BAM8* (these genes have changed transcripts in antisense-*ACLA* plants compared to that of WT at these three time points, data not shown). We are interested both in genes that are differentially expressed across all three time points, and also at single time points.

To identify other genes whose transcript levels are altered in antisense-*ACLA* plants relative to WT, we compared the global gene expression over the 3 time points between antisense-*ACLA* and WT. Sixty nine genes are significant in the "Genotype" comparison when FDR is controlled at the level of 0.2 using the method of Storey and Tibshirani (2003). The 69 genes can be grouped into 3 clusters by K-medoids clustering method (see Materials and Methods) (Figure 9A and Appendix E, Table S2). Each cluster has quite a few genes that are annotated as "expressed protein" and "functions unknown". Genes in Cluster 1 have increased expression in the antisense-*ACLA.* Among them, are several genes involved in regulation of transcription, and also in protein posttranslational modification, cell wall metabolism, secondary metabolism (e.g., flavonoids), lipid metabolism (e.g., steroids), hormone metabolism (e.g., auxin and ethylene) and stress response (Appendix E, Table S2). Genes in Cluster 2 and 3 have decreased expression in the antisense-*ACLA,* including genes involved in stress response, regulation of transcription, cell wall metabolism, secondary metabolism (e.g., flavonoids), hormone metabolism (e.g., auxin) and protein metabolism.

Sixty one genes are significant in the "Genotype*Time" comparison when FDR was restricted at the level of 0.45 using the method of Storey and Tibshirani (2003). They can be grouped into 4 clusters by K-medoids clustering method (see Materials and Methods) (Figure 9B and Appendix E, Table S3). This "Genotype*Time" comparison selects expression patterns that diverse compared to the "Genotype" comparison. Each cluster has several genes that are annotated as "expressed protein" and "functions unknown". Though this estimate of FDR is very high, there are still genes showing different expression patterns between the antisense-*ACLA* mutant and WT in Cluster 1, 2 and 4. These clusters contain genes whose transcripts are altered in opposite direction between antisense-*ACLA* and WT, and are involved in protein posttranslational modification, cell wall metabolism, hormone metabolism and stress response.

It is not yet understood how reduced cytosolic ACL activity redirects the carbon flow into starch. We checked the expression of genes in all pathways in the MetNetDB (Wurtele et al., 2003) (http://www.metnetdb.org/MetNet_db.htm). If at any time point, the error bars of gene expression are not overlapped between the genotypes, this gene is considered as expression altered at that time point. Quite a few pathways have genes that are up-regulated

or down-regulated in the antisense-*ACLA* (Table 2), e.g., fatty acid elongation, cutin biosynthesis, mevalonate pathway, biotin biosynthesis, brassinosteroid biosynthesis, and chlorophyll biosynthesis, have genes that are only up-regulated in the antisense-*ACLA*; while brassinosteroids signaling, abscisic acid/IAA/jasmonic acid biosynthesis, several pathways of sulfur containing amino acid biosynthesis (e.g., homocysteine and cysteine interconversion, cysteine biosynthesis), and APX1 signal transduction pathway have genes that are only down-regulated in the antisense-*ACLA*. And also, genes that have altered transcripts in antisense-*ACLA* in ascorbate glutathione pathway (cytosol), ascorbate biosynthesis, ascorbate glutathione pathway (mitochondria), ascorbate glutathione pathway (plastid stroma) and ascorbate glutathione cycle have decreased expression except superoxide dismutase (in cytosol and chloroplast), ferritin 1 (in chloroplast) and chorismate mutase (in cytosol) (data not shown). These pathways help to keep active oxygen under control.

Some processes have genes with transcripts altered in both directions. Particularly, pathways involved in carbohydrate metabolism, especially starch metabolism, trehalose metabolism, and cell wall metabolism, have genes with fluctuated expression alteration in the antisense-*ACLA*. To determine how decreased cytosolic ACL activity affects both starch metabolism genes, and other genes implicated in the carbohydrate network, we focused on pathways that are sensitive to carbon signaling and to changes in metabolic flux, and that direct the net accumulation of starch and sugars. Thirty five different genes in trehalose, and cell wall metabolism showed altered expression over the three time points: 1 h in the light, 0.5 h and 4 h in the dark (Table 3). Genes in Table 3 have no less than 2-fold change between antisense-*ACLA* and WT for at least one time point, and the error bars are not overlapped between genotypes at that time point. Functional categorization of these genes is based on information from MetNetDB (Wurtele et al., 2003), MapMan (Thimm et al., 2004) and from TAIR and GO databases using AtGeneSearch (http://www.metnetdb.org/MetNet_atGeneSearch.htm).

Expression profiles of most of the 29 starch metabolic genes are similar between WT and antisense-*ACLA* mutants, e.g., *GBSS* (Figure 10A). However, four starch metabolic genes show different levels of expression in antisense-*ACLA* compared to WT: *AAM1* (At1g76130), *AAM2* (At4g25000), *BAM1* (At4g15210) and *BAM3* (At3g23920) (Figure

10B). Except *BAM3*, the other three genes have reduced transcripts in antisense-*ACLA*. But only *AAM2* and *BAM1* have more than 2-fold change in expression between antisense-*ACLA* and WT. Of the genes related to starch metabolism, four genes: *APL3* (At2g21590), *APL4* (At4g39210), *GPT1* (At5g54800) and *GPT2* (At1g61800) (G-6-P/P transporter), all these show decreased RNA in antisense-*ACLA* than in WT.

Though both ACLA and ACLB proteins levels are reduced in the antisense-*ACLA* plants (Fatland et al., 2005), there are no clear RNA accumulation changes of ACL genes except *ACLA2* (At1g60810) (Figure 10C). Predominant *ACLA2* transcript is reduced in antisense-*ACLA* plants at all three time points; this might indicate that ACL activity is affected at the post-transcriptional or translational level.

Trehalose metabolism

Trehalose, a non-reducing sugar composed of two glucose molecules, is synthesized by trehalose-6-P synthase (TPS) and trehalose-6-P phosphatase (TPP) in two catalytic steps. At specific times during the diurnal cycle, there is more than a 2-fold difference in accumulation of *TPP*s (three isoforms) transcripts between antisense-*ACLA* and WT.

Cell Wall

Thirty-two genes implicated in cell wall metabolism are differentially regulated in the antisense-*ACLA* mutant. Both up- and down-regulated genes were involved with the synthesis, modification, and degradation of the cell wall.

Three cell wall synthesis genes whose mRNAs are up-regulated in antisense-*ACLA*: one UDP-galactose-4-epimerase, and two cellulose synthase-like proteins (Csl) which are putative glycosyltransferases belonging to the cellulose synthase superfamily (Williamson et al., 2002).

Hemicellulose synthesis genes whose transcripts are down-regulated in the antisense-*ACLA* mutant are mannose-6-phosphate isomerase 1, rhamnosyltransferase/ UDP-glucoronosyl and –glucosyl transferase (hemicellulose), and UDP-glucose/galactose 4-epimerase-like protein (NAD-dependent).

Up-regulated cell wall modification genes included two beta-expansins, one alpha-expansin, four endo-xyloglucan transferases (or xyloglucan endotransglycosylases. In addition, a pectin methylesterase (PME), and two invertase/pectin methylesterase inhibitors

are up-regulated. During pectin modification and degradation, pectin esterases and pectin acetylesterase are required for the removal of carboxyl or hydroxyl groups. Invertase/ pectin methylesterase inhibitors are extracellular proteins that specifically target either invertases or PMEs, but are grouped into a large gene family based on sequence similarity (Hothorn et al., 2004).

Six genes involved with cell wall modification are down-regulated in antisense-*ACLA*; one endo-xyloglucan transferase, one invertase/pectin methylesterase inhibitor, three pectin esterases, and one pectin acetylesterase.

Four genes implicated in the degradation of cell walls in plants are up-regulated. These include one polygalacturonase (pectinase), one beta-xylosidase, one (1,4)-beta-mannan endohydrolase, and a pectate lyase. Polygalacturonases are endo- or exo-acting enzymes that degrade de-esterified pectin through hydrolytic cleaving of chain residues (Brummell and Harpster, 2001). Beta- and alpha-xylosidases remove xylose residues from side chains and backbone of the non-reducing end of xyloglucan oligosaccharides, respectively (Sampedro et al., 2001; Minic et al., 2004).

Six cell wall degradation genes are down-regulated in antisense-*ACLA*: four xyloglucan endo-1,4-beta-D-glucanases, are members of the large EGase family (Nicol et al., 1998); one beta-xylosidase, and one pectate lyase. The polysaccharide-rich cell walls of plants are degraded by endopolygalacturonases produced by fungi. In order to control this damage, plants produce cell wall-localized polygalacturonase inhibiting proteins (PGIPs) (De Lorenzo et al., 2001).

Genes from the cell wall, trehalose, starch, sucrose, and acetyl-CoA metabolism pathways can be placed in a small network (Figure 11A). Figure 11B includes some selected pathways to show the gene expression change.

DISCUSSION

Co-regulation of genes encoding constituents of the same metabolic pathway or protein complex might be discernable if clusters of genes with related functions emerge from pattern correlation analyses of large data sets of transcriptomics data spanning many biotic and abiotic conditions (Eisen et al., 1998; Segal et al., 2003; Stuart et al., 2003).

Of the three distinct groups of co-expressed genes within the starch metabolism network, the genes whose expression peaks in the dark, also described as occurring under LD conditions by Smith et al. (2004), appeared to be involved in starch degradation. These include six genes (*DPE2*, *GWD1*, *GWD3*, *PHS2*, *SEX4*, and *ISA3*) already implicated in starch degradation (Smith et al., 2005; Niittylä et al., 2006). The other two genes co-expressed (in starch metabolic pathway) (*BAM2*, *BE3*) have no previously identified function; they might also take part in starch degradation. The second group of co-expressed starch metabolic genes that peak in the light, containing *ISA1*, *ISA2*, *PHS1* and *DPE1*, appeared to be involved in starch synthesis. The function of the remaining group of co-expressed genes is unclear. Interestingly, only five genes, all of which are identified by sequence homology only as "putative" beta- and alpha-amylases (*BAM9*, *BAM3*, *BAM6*, *BAM1*, and *AAM2*), are not highly correlated to any other known starch metabolic gene, nor to each other.

Light triggers complex reactions to mediate plant growth and development. The seven clusters of co-expressed genes in WT over the SD diurnal cycle may reflect co-regulation of processes over the diurnal cycle. An example is the co-expression of the set of genes (*COR*, *ERD*, *CCR*, and *AtGRP*) which have been reported to be cold regulated (Gilmour et al., 2000). Eight of these cold-regulated genes also respond to a diurnal cycle with peaks in expression at the beginning of the dark. The plants in this experiment were grown at a constant temperature, thus the change in expression was not triggered by a change in temperature. Four of the five genes in the DREB1 and DREB2 transcriptional activator families, also have peaks at the beginning of the dark. Only the clusters (Cluster 5 and 6, Figure 3) with gene expression peaks at the beginning of dark contain the "cold responsive" genes. And these clusters are also rich in other genes involved in stress response. Another possibility is that this mechanism provides the cell with appropriate levels of RNA to respond to changes in environmental conditions, and finer controls in protein level are modulated post-transcriptionally.

In antisense-*ACLA* plants, the higher starch content may indicate an imbalance between starch biosynthesis and mobilization, reflecting stimulated starch synthesis, reduced starch degradation, or both. Either would result in more accumulation of starch. In the

mutant, several genes, which are implicated in starch synthesis, specifically, *APL3*, *APL4*, *GPT1*, *GPT2* and three *TPP*s, have up-regulated transcripts in the mutant relative to WT. Applied sucrose induces *APL3* and *APL4* transcription in leaves (Crevillén et al., 2005). Trehalose and sucrose have been postulated to induce starch synthesis in Arabidopsis source tissue via an increase in ADP-glucose pyrophosphorylase activity and *APL3* expression (Wingler et al., 2000; Rook et al., 2001). Glucose 6-phosphate/phosphate translocator has been proposed to import glucose-6-P into plastids to be used as a precursor for starch in part because *GPT1* and *GPT2* (At1g61800) could rescue the *pgi1* mutant from low starch in leaf (Niewiadomski et al., 2005). In contrast, *BAM1*, *AAM1* and *AAM2*, have down-regulated transcripts in the mutant relative to WT. These genes have amylase-like sequence, however, the sub-cellular localization and role of AAM1 and AAM2 remains unclear (Yu et al., 2005); and BAM1 has not been functionally identified. We propose that up-regulated transcripts of genes implicated in starch synthesis, and possibly down-regulated transcripts of potential starch degradative genes, contribute to the higher starch accumulation in the antisense-*ACLA* mutant, relative to WT.

The reduced ACL expression presumably results in an increase of starch indirectly, through a potentially complex metabolic and regulatory network. Some of this network may be transcriptionally controlled, and thus potentially revealed by the microarray data. For the majority of genes, transcript accumulation patterns are unaffected by the ACL mutation. For example, the diurnal RNA accumulation patterns of circadian rhythm genes are similar between antisense-*ACLA* mutant and WT, indicating that the antisense-*ACLA* mutation does not alter transcript accumulation of circadian rhythm genes.

However, the accumulation of a subset of transcripts is altered in the antisense-*ACLA* mutant. Several pathways involved in fatty acid elongation, secondary metabolite or hormone biosynthesis, have genes that are up-regulated in antisense-*ACLA*. These pathways require substrate derived from acetyl-CoA (Fatland et al., 2005). In antisense-*ACLA* lines, the impact of diminished ACL activity could reduce the acetyl-CoA pool. A diminished supply of substrate might stimulate the expression of genes in selected biosynthetic pathways.

In the mutant, two stress-induced hormone pathways have down-regulated gene expression, i.e., synthesis of JA and ABA. Also, cytokinin degradation genes are down-

regulated. In addition, of the 13 genes involved in the APX1 signal transduction pathways, five have reduced transcript accumulation in the antisense-*ACLA* mutant. One of these is *APX1* (At1g07890, ascorbate peroxidase) which encodes the cytosolic antioxidant protein, relieving oxidative damage in the chloroplast (Davletova et al., 2005). The down-regulation of the APX pathway(s), might indicate an oxidative imbalance in the antisense-*ACLA* plants. Other genes involved in oxidative stress are also down-regulated, these include peptide methionine sulfoxide reductase2 (PMSR2, At5g07460), which has been reported to repair cellular oxidative damage in long night (Bechtold et al., 2004). Starch synthesis is thought to be regulated by redox-mediated modification of AGPase (Hendriks et al., 2003). One possibility–garnered from the transcription data, is that in the oxidatively-stressed-antisense-*ACLA* plants, an altered redox would increase AGPase activity and maybe also contribute to the increase in starch content in the antisense-*ACLA* plants.

Trehalose metabolism plays a role in stress protection, carbohydrate storage, and glycolytic regulation in many microbes (Goddijn and van Dun, 1999). Its role in plants has not been well-elucidated, but trehalose metabolism has been implicated in aspects of stress response (Avonce et al., 2004) and in central carbon metabolism (Eastmond et al., 2003; Wingler et al., 2000). The up-regulated *TPP* transcripts in the antisense-*ACLA* plants may reflect an interconnection between oxidative stress and carbon metabolism.

Thirty five genes involved in cell wall metabolism, including synthesis, modification and degradation, are either up-regulated or down-regulated in the antisense-*ACLA* plants relative to WT. The complex modifications of cell wall components are not fully understood. The expression of six genes (and gene candidates) synthesizing hemicellulose constituents is altered. The alteration in transcripts mostly takes place at half an hour in dark. Another 16 genes related to cell wall modification with altered transcripts include those might take a role in cell growth. Most of the alteration in transcripts happens at 4 h in the dark. These genes are expansins and endo-xyloglucan transferases, which are thought to cooperate during cell growth for cell wall assembly, loosening, and restructuring (Peña et al., 2004). The 10 genes with altered transcripts, in cell wall degradation, mostly occur at 1 h in light. These genes are thought to function in plant defense. The middle lamella of plant cell wall are degraded by pectate lyase secreted by plant pathogens; 27 pectin-lyase like genes are present in

Arabidopsis (Marin-Rodriguez et al., 2002). *PGIP2* transcripts accumulate in response to wounding and pathogen infection, as well as, defense-related signals such as methyl jasmonate and salicylic acid (Ferrari et al., 2003).

Acetyl-CoA plays a critical role for many biochemical pathways essential to plant growth and development (Fatland et al., 2005). Our working hypothesis is that decreased cytosolic ACL results in insufficient cytosolic acetyl-CoA substrate for generation of the isoprenoids and polyketides derived from this molecule. This alteration in metabolites could affect the plant signaling. Specifically, down-regulated APX1 signal transduction pathway might induce oxidative stress. It is possible that redox-regulation of starch metabolism result in a carbon partitioning into starch. We envision these processes could affect and regulate each other. They represent testable hypotheses that could elucidate metabolic interconnections between starch accumulation and acetyl-CoA metabolism, two disparate metabolic processes.

MATERIALS AND METHODS

Plant Materials, Growth Conditions, Experimental Design

Plants used in this study were wild-type *Arabidopsis thaliana* (ecotype Columbia) and transgenic *A. thaliana* in the Columbia background containing antisense *ACLA* behind the constitutive CaMV 35S promoter (Fatland, 2002). Seeds were surface sterilized in a 50% bleach/ 0.02% Triton X-100 solution for 7 min and rinsed in sterile distilled H_2O three times. Antisense-*ACLA* and wild-type seeds were germinated on a solid growth medium, with and without 30 μg mL^{-1} kanamycin, respectively. The growth medium was buffered with 2.56 mM MES at pH 5.7 and contained 4.3 g/L Murashige and Skoog salts (GibcoBRL, Life Technologies, Rockville, MD), 1% sucrose, 1X B5 vitamins (100 μg/mL myo-inositol, 1 μg/mL pyridoxine hydrochloride, 1 μg/mL nicotinic acid, 10 μg/mL thiamine hydrochloride), and 0.6% (w/v) agargel (Sigma, St. Louis, MO). The seeds were placed in a growth chamber (73% RH at 22.2 ± 1.1°C) under a short day (SD) photoperiod (8 h light /16 h dark) from fluorescent lamps (129 ± 16 $\mu M \cdot m^{-2} \cdot s^{-1}$ PAR). After 21 d, seedlings were transplanted to a sterile potting medium (LC1 Sunshine Mix, Sun Gro, Horticulture, Inc., Bellevue, WA) and fertilized with 1X Nutriculture Grower's Special Blend 21-8-18^{PLUS}

(Plant Marvel Labs, Chicago, Il). Only antisense-*ACLA* seedlings exhibiting features characteristic of the antisense phenotype were transplanted (Fatland, 2002). Plants were arranged in 16 randomized blocks (flats) using PROC PLAN of SAS (SAS Institute, Cary, NC) and returned to the same growth chamber and environmental conditions as used for germinating seeds. Each flat consists of plants of the same genotype (8 flats are of WT and 8 flats are of antisense-*ACL*). There were 21 pots in one flat with 2 plants / pot. Rosette leaves (#5 to #8) from 2 plants in one pot from each flat were harvested from six-week-old seedlings at 0, 0.5, 1, 4, 6, 8, 8.5, 9, 12, 14, 16, and 20 h during the SD photoperiod. Time points 0, 0.5, 1, 4, and 6 h were collected during the light cycle, and time points 8, 8.5, 9, 12, 14, 16, and 20 h were collected during the dark cycle under a green light. For each time point, there were 2 samples (of 2 genotypes) harvested. Each sample consisted of leaves from sixteen plants. The entire process was completed on two separate occasions to obtain two replications for all genotype and time combinations. This can be viewed as a split-plot experimental design with genotype as the whole-plot factor and time as the split-plot factor. Samples were frozen in liquid N_2 immediately after harvest and stored at -80^{o}C for RNA and protein extraction.

Starch Quantification

Plants were grown under the same conditions as described above. At 6 h into the light phase of the photoperiod, samples of 250 to 500 mg fresh weight rosette leaves were harvested from randomly selected plants. Three samples, consisting of leaves from five plants, were collected per genotype. Samples were frozen in liquid N_2 immediately after harvest and stored at -80°C. Starch content of leaf samples was determined as described by Zeeman et al. (1998) with slight modification as follows. Leaf samples were boiled in 35 mL 80% (v/v) ethanol for 30 min or until completely de-pigmented. Insoluble material was homogenized in a Kontes tissue grinder and extracted again with 35 mL 80% (v/v) ethanol. After homogenizing the insoluble material a second time, it was resuspended in water and boiled for 30 min. Total α-glucan polymer content (soluble and insoluble) was quantified with a commercial assay kit (catalog no. E0207748, R-BioPharm, Darmstadt, Germany) that

measures glucose released following digestion with amyloglucosidase. T test was performed by using JMP software (SAS Institute).

Northern Blot Analysis

Total RNA was extracted from leaves with GibcoBRL TRIzol reagent (Life Technologies) according to the protocol provided with the reagent. RNA concentrations were measured using a Genesys spectrophotometer (Spectronic Instruments, Rochester, NY). RNA samples (10 μg) were denatured for 15 min at 65°C in loading buffer and resolved in a 1.0% denaturing agarose gel by using 0.67% (220 mM) formaldehyde in the gel and in the 1X MOPS (40 mM MOPS [3-[N-morpholino]propane sulfonic acid]; 1 mM EDTA; 10 mM sodium acetate, pH 7.0) running buffer. RNA was transferred onto nylon membranes with 10X SSC. RNA markers (0.24-9.5 kb) were used to estimate size of transcripts. Pre-hybridization and hybridization of the membranes were done at 42°C in 50% (w/v) formamide, 6.7X SSC, 3.3X Denhardt's solution, 0.4% (w/v) SDS, 25 mM sodium phosphate buffer (pH 7), and 0.12 μg μL^{-1} salmon sperm DNA. cDNA probes were labeled with [^{32}P]dCTP by random priming using GibcoBRL reagents and protocol (Life Technologies). After hybridization, the membrane was rinsed for 5 min at room temperature in 2X SSC and 0.1% SDS. Two washes, 15 min each, were done at 65°C in 1X SSC and 0.1% SDS. Washed membranes were exposed to a phosphor screen (Molecular Dynamics, Sunnyvale, CA) for 4 d, scanned with a Storm 840 PhosphorImager (Molecular Dynamics), and visualized with ImageQuant software (Molecular Dynamics). To confirm equal loading of RNA per lane, membranes were stripped of the former probe and rehybridized with a wheat *18S* rDNA probe.

GeneChip Analysis of SD Experiments

Total RNA was extracted as described for northern blot analysis. Synthesis of labeled cRNAs, hybridization with Arabidopsis GeneChips (Affymetrix, Santa Clara, CA), and scanning of the probe array were performed at the Iowa State University Gene Chip Facility (Ames, IA). Hybridization intensities were collected by an Agilent GeneArray Scanner (Agilent Technologies, Palo Alto, CA). For data analysis, relative expression intensities were

generated and normalized with Microarray Suite (MAS) 5.0 software (Affymetrix) in the form of signal values. The signal value, a robust weighted mean of probe fluorescence for a probe set (corrected for nonspecific signal by subtracting mismatch probe values), was calculated using the One-step Tukey's Biweight Estimate. Global scaling (Affymetrix) was used to normalize the data by adjusting the mean expression level of each chip to a value of 500. Scatter plot analysis of the two replicates were linear indicating similarity to each other (data not shown); we thus used the mean value of expression (normalized intensity) for the graphs in Figure 4 and 5. FCModeler (Iowa State University, Ames, IA) was used to model and visualize metabolic networks.

Clustering

The data from each chip was normalized by log (signal+1) transformed. For each gene, we fit a mixed linear model with fixed effects for time and random effects for replicate to identify genes that exhibit significant change in expression across 11 time points for WT; we fit another mixed linear model with fixed effects for time and genotype and random effects for replicate to identify genes that exhibit significant change in expression over the 3 time points per genotype and 2 genotypes surveyed respectively. When testing for differences over time, we used the more conservative Benjamini and Hochberg method (1995) to estimate FDR because a large number of genes exhibited significant differences over time. We used the method of Storey and Tibshirani (2003) to estimate FDR to gain more power for detecting differences when comparing the WT and mutant genotypes. Thresholds for significance based on q-values were selected to maintain false discovery below desired rates. After matrix of estimated expression means was obtained, the matrix of means was standardized so that Euclidean distance would be related to 1-correlation distance. K-medoids clustering was conducted with the cluster number that Gap analysis recommended. The Gap statistic method (Tibshirani et al., 2001) calculates the number of clusters by comparing the degree of clustering of the data to that of a simulated uniform (cluster-free) reference distribution.

Correlation Analysis

Arabidopsis Affymetrix ATH1 data from 963 chips were obtained from Nottingham Arabidopsis Stock Centre microarray database (http://affymetrix.arabidopsis.info/). The data represent various experiments, including development, stress, mutant, and other (http://affymetrix.arabidopsis.info/narrays/experimentbrowse.pl) studies. All chips were individually scaled to a common mean=100, excluding top and bottom 2% signal intensities, using MAS 5.0 (Affymetrix). We averaged the expression values on chips coming from biological replicates after assessing their replicability by comparing the interquantile range of each data set. Pair-wise Pearson correlation coefficients were computed for 58 genes of the starch metabolism pathway across all probes on the Arabidopsis Affymetrix ATH1 chips by using our custom software, MetaOmGraph (http://www.metnetdb.org/MetNet_MetaOmGraph.htm). From among the 22K genes represented on the chips, genes correlated above a specified cut-off to a gene of interest were identified and visualized with MetaOmGraph. Correlation data, generated in R software (http://www.r-project.org/index.html) (R Development Core Team, 2004), for starch metabolism genes were clustered by columns and rows in CIMMaker (http://discover.nci.nih.gov/nature2000/tools/cimmaker.jsp) and visualized in Matrix2png (http://microarray.cpmc.columbia.edu/matrix2png/) (Pavlidis and Noble, 2003).

Activity Gel Analyses of Starch Synthesizing and Modifying Enzymes

Zymogram analyses of starch synthesizing and modifying enzymes were conducted as described by Cao et al. (1999) and Dinges et al. (2001), respectively, except leaf samples were ground in 50 mM Tris acetate, pH 7.5 and 10 mM DTT (5 mL/g fresh weight), and repeated three times. Protein samples (200 μg) were separated under non-denaturing conditions in native 7% polyacrylamide gels, with and without 0.3% (w/v) glycogen, for 4.5 h at 4ºC at 75 mA in a Protean II cell (Bio-Rad Laboratories, Hercules, CA). Following electrophoresis, the gel containing glycogen was incubated in reaction medium (100 mM Bicine, pH 8.0, 0.5 M citrate, 25 mM potassium acetate, 0.5 mg/mL BSA, 5 mM ADP-Glc, 5 mM 2-mercaptoethanol, and 20 mg/mL glycogen) at room temperature for 36 h to facilitate starch synthase activity. For the starch modification assay, proteins were transferred to a polyacrylamide gel containing 0.3% (w/v) starch by electroblotting. Iodine stain (0.2% iodine

and 2% potassium iodide in 10 mM HCl) was used to detect enzyme activity in the gels. Stained gels were scanned immediately with a GS-800 Calibrated Densiometer (Bio-Rad), and visualized with MagicScan software (UMAX Technologies, Fremont, CA).

ACKNOWLEDGMENTS

We would like to thank Tracie Bierwagen, Dr. Christophe Colleoni, and Dr. Jennifer Walker-Daniels for helpful discussions regarding starch metabolism. We are particularly grateful to Dr. Pan Du and Jie Li for valuable discussion of FCModeler and MetNetDB. The LHCB3 cDNA was purchased from RIKEN.

LITERATURE CITED

Avonce N, Leyman B, Mascorro-Gallardo JO, Van Dijck P, Thevelein JM, Iturriaga G (2004) The Arabidopsis trehalose-6-P synthase *AtTPS1* gene is a regulator of glucose, abscisic acid, and stress signaling. Plant Physiol **136:** 3649–3659

Baier M, Hemmann G, Holman R, Corke F, Card R, Smith C, Rook F, Bevan MW (2004) Characterization of mutants in Arabidopsis showing increased sugar-specific gene expression, growth, and developmental responses. Plant Physiol **134:** 81–91

Baroja-Fernández E, Muñoz FJ, Zandueta-Criado A, Morán-Zorzano MT, Viale AM, Alonso-Casajús N, Pozueta-Romero J (2004) Most of ADP-glucose linked to starch biosynthesis occurs outside the chloroplast in source leaves. PNAS **101:** 13080–13085

Bechtold U, Murphy DJ, Mullineaux PM (2004) Arabidopsis peptide methionine sulfoxide reductase2 prevents cellular oxidative damage in long nights. Plant Cell **16:** 908–919

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J Roy Statist Soc Ser B **57:** 289–300

Bogdanove AJ, Martin GB (2000) AvrPto-dependent Pto-interacting proteins and AvrPto-interacting proteins in tomato. PNAS **97:** 8836–8840

Bowien B, Kusian B (2002) Genetics and control of CO_2 assimilation in the chemoautotroph *Ralstonia eutropha*. Arch Microbiol **178:** 85–93

Bowman J (1994) Arabidopsis: an atlas of morphology and development. Springer-Verlag, New York.

Brummell DA, Harpster MH (2001) Cell wall metabolism in fruit softening and quality and its manipulation in transgenic plants. Plant Mol Biol **47:** 311–340

Cao H, Imparl-Radosevich J, Guan H, Keeling PL, James MG, Myers AM (1999) Identification of the soluble starch synthase activities of maize endosperm. Plant Physiol **120:** 205–215

Chia T, Thorneycroft D, Chapple A, Messerli G, Chen J, Zeeman SC, Smith SM, Smith AM (2004) A cytosolic glucosyltransferase is required for conversion of starch to sucrose in Arabidopsis leaves at night. Plant J **37:** 853–863

Coruzzi G, Bush DR (2001) Nitrogen and carbon nutrient and metabolite signaling in plants. Plant Physiol **125:** 61–64

Craigon DJ, James N, Okyere J, Higgins J, Jotham J, May S (2004) NASCArrays: a repository for microarray data generated by NASC's transcriptomics service. Nucleic Acids Res **32:** D575–D577

Crevillén P, Ballicora MA, Mérida Á, Preiss J, Romero JM (2003) The different large subunit isoforms of Arabidopsis thaliana adp-glucose pyrophosphorylase confer distinct kinetic and regulatory properties to the heterotetrameric enzyme. J Biol Chem, **278:** 28508–28515

Crevillén P, Ventriglia T, Pinto F, Orea A, Mérida Á, Romero JM (2005) Differential pattern of expression and sugar regulation of Arabidopsis thaliana ADP-glucose pyrophosphorylase-encoding genes. J Biol Chem **280:** 8143–8149

Cross JM, Clancy M, Shaw JR, Greene TW, Schmidt RR, Okita TW, Hannah LC (2004) Both subunits of ADP-glucose pyrophosphorylase are regulatory. Plant Physiol **135:** 137–144

Davletova S, Rizhsky L, Liang H, Shengqiang Z, Oliver DJ, Coutu J, Shulaev V, Schlauch K, Mittler R (2005) Cytosolic ascorbate peroxidase 1 is a central component of the reactive oxygen gene network of Arabidopsis. Plant Cell **17:** 268–281

De Lorenzo G, D'Ovidio R, Cervone F (2001) The role of polygalacturonase-inhibiting proteins (PGIPs) in defense against pathogenic fungi. Annu Rev Phytopathol **39:** 313–335

Dinges JR, Colleoni C, Myers AM, James MG (2001) Molecular structure of three mutations at the maize *sugary1* locus and their allele-specific phenotypic effects. Plant Physiol **125:** 1406–1418

Eastmond PJ, Li Y, Graham IA (2003) Is trehalose-6-phosphate a regulator of sugar metabolism in plants? J Exp Bot **54:** 533–537

Eisen MB, Spellman PT, Brown PO, Botstein D (1998) Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci USA **95:** 14863–14868

Fatland BL, Ke J, Anderson MD, Mentzen WI, Cui LW, Allred CC, Johnston JL, Nikolau BJ, Wurtele ES (2002) Molecular characterization of a heteromeric ATP-citrate lyase that generates cytosolic acetyl-coenzyme A in Arabidopsis. Plant Physiol **130:** 740–756

Fatland BL, Nikolau BJ, Wurtele ES (2005) Reverse genetic characterization of cytosolic acetyl-CoA generation by ATP-citrate lyase in Arabidopsis. Plant Cell **17:** 182–203

Ferrari S, Vairo D, Ausubel FM, Cervone F, De Lorenzo G (2003) Tandemly duplicated Arabidopsis genes that encode polygalacturonase-inhibiting proteins are regulated coordinately by different signal transduction pathways in response to fungal infection. Plant Cell **15:** 93–106

Fordham-Skelton AP, Chilley P, Lumbreras V, Reignoux S, Fenton TR, Dahm CC, Pages M, Gatehouse JA (2002) A novel higher plant protein tyrosine phosphatase interacts with SNF1-related protein kinases via a KIS (kinase interaction sequence) domain. Plant J **29:** 705–715

Fox TC, Geiger DR (1984) Effects of decreased net carbon exchange on carbohydrate metabolism in sugar beet source leaves**.** Plant Physiol **76:** 763–768

Geiger DR, Servaites JC (1994) Diurnal regulation of photosynthetic carbon metabolism in C3 plants. Annu Rev Plant Physiol Plant Mol Biol **45:** 235–256

Gibon Y, Blasing OE, Palacios-Rojas N, Pankovic D, Hendriks JH, Fisahn J, Hohne M, Gunther M, Stitt M (2004) Adjustment of diurnal starch turnover to short days: depletion of sugar during the night leads to a temporary inhibition of carbohydrate utilization, accumulation of sugars and post-translational activation of ADP-glucose pyrophosphorylase in the following light period. Plant J **39:** 847–62

Gilmour SJ, Sebolt AM, Salazar MP, Everard JD, Thomashow MF (2000) Overexpression of the Arabidopsis CBF3 transcriptional activator mimics multiple biochemical changes associated with cold acclimation. Plant Physiol **124:** 1854–1865

Goddijn OJM, van Dun K (1999) Trehalose metabolism in plants. Trends Plant Sci **4:** 315–319

Halford NG, Hardie DG (1998) SNF1-related protein kinases: global regulators of carbon metabolism in plants? Plant Mol Biol **37:** 735–748

Harmer SL, Hogenesch JB, Straume M, Chang HS, Han B, Zhu T, Wang X, Kreps JA, Kay SA (2000) Orchestrated transcription of key pathways in Arabidopsis by the circadian clock. Science **290:** 2110–2113

Hayashi NR, Terazono K, Kodama T, Igarashi Y (2000) Structure of ribulose 1,5-bisphosphate carboxylase/oxygenase gene cluster from a thermophilic hydrogen-oxidizing bacterium, Hydrogenophilus thermoluteolus, and phylogeny of the fructose 1,6-bisphosphate aldolase encoded by cbbA in the cluster. Biosci Biotechnol Biochem **64:** 61–71

Hendriks JHM, Kolbe A, Gibon Y, Stitt M, Geigenberger P (2003) ADP-glucose pyrophosphorylase is activated by posttranslational redox-modification in response to light and to sugars in leaves of Arabidopsis and other plant species. Plant Physiol **133:** 838–849

Hothorn M, Wolf S, Aloy P, Greiner S, Scheffzek K (2004) Structural insights in the target specificity of plant invertase and pectin methylesterase inhibitory proteins. Plant Cell **16:** 3437–3447

Hussain H, Mant A, Seale R, Zeeman S, Hinchliffe E, Edwards A, Hylton C, Bornemann S, Smith AM, Martin C, Bustos R (2003) Three isoforms of isoamylase contribute different catalytic properties for the debranching of potato glucans. Plant Cell **15:** 133–149

Kakefuda G, Duke SH (1984) Electrophoretic transfer as a technique for the detection and identification of plant amylolytic enzymes in polyacrylamide gels. Plant Physiol **75:** 278–280

Koch K (1996) Carbohydrate-modulated gene expression in plants. Annu Rev Plant Physiol Plant Mol Biol **47:** 509–540

Liu Q, Kasuga M, Sakuma Y, Abe H, Miura S, Yamaguchi-Shinozaki K, Shinozaki K (1998) Two transcription factors, DREB1 and DREB2, with an EREBP/AP2 DNA binding domain separate two cellular signal transduction pathways in drought- and low-temperature-responsive gene expression, respectively, in Arabidopsis. Plant Cell **10:** 1391–1406

Lu Y, Gehan JP, Sharkey TD (2005) Daylength and circadian effects on starch degradation and maltose metabolism. Plant Physiol **138:** 2280–2291

Lu Y, Sharkey TD (2004) The role of amylomaltase in maltose metabolism in the cytosol of photosynthetic cells. Planta **218:** 466–473

Marin-Rodriguez MC, Orchard J, Seymour GB (2002) Pectate lyases, cell wall degradation and fruit softening. J Exp Bot **53:** 2115–2119

Matt P, Schurr U, Klein D, Krapp A, Stitt M (1998) Growth of tobacco in short-day conditions leads to high starch, low sugars, altered diurnal changes in the Nia transcript and low nitrate reductase activity, and inhibition of amino acid synthesis. Planta **207:** 27–41

Mentzen IW (2006) From pathway to regulon in Arabidopsis. Ph.D. dissertation. Iowa State University, Ames, IA

Minic Z, Rihouey C, Do CT, Lerouge P, Jouanin L (2004) Purification and characterization of enzymes exhibiting β-D-xylosidase activities in stem tissues of Arabidopsis. Plant Physiol **135:** 867–878

Myers AM, Morell MK, James MG, Ball SG (2000) Recent progress toward understanding biosynthesis of the amylopectin crystal. Plant Physiol **122:** 989–997

Neff MM, Fankhauser C, Chory J (2000) Light: an indicator of time and place. Genes Dev **14:** 257–271

Nicol F, His I, Jauneau A, Vernhettes S, Canut H, Höfte H (1998) A plasma membrane-bound putative endo-1,4-β-D-glucanse is required for normal wall assembly and cell elongation in Arabidopsis. EMBO J **17:** 5563–5576

Niewiadomski P, Knappe S, Geimer S, Fischer K, Schulz B, Unte US, Rosso MG, Ache P, Flugge UI, Schneider A (2005) The Arabidopsis plastidic glucose 6-phosphate/phosphate translocator GPT1 is essential for pollen maturation and embryo sac development. Plant Cell **17:** 760–775

Niittylä T, Comparot-Moss S, Lue WL, Messerli G, Trevisan M, Seymour MDJ, Gatehouse JA, Villadsen D, Smith SM, Chen J, Zeeman SC, Smith AM (2006) Similar protein phosphatases control starch metabolism in plants and glycogen metabolism in mammals. J Biol Chem **281:** 11815–11828

Niittylä T, Messerli G, Trevisan M, Chen J, Smith AM, Zeeman SC (2004) A previously unknown maltose transporter essential for starch degradation in leaves. Science **303:** 87–89

Pavlidis P, Noble WS (2003) Matrix2png: A Utility for Visualizing Matrix Data. Bioinformatics **19:** 295–296

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

R Development Core Team (2004) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL: http://www.r-project.org

Riemenschneider A, Wegele R, Schmidt A, Papenbrock J (2005) Isolation and characterization of a D-cysteine desulfhydrase protein from Arabidopsis thaliana. FEBS J **272:** 1291–1304

Rook F, Corke F, Card R, Munz G, Smith C, Bevan MW (2001) Impaired sucrose-induction mutants reveal the modulation of sugar-induced starch biosynthetic gene expression by abscisic acid signalling. Plant J **26:** 421–433

Sampedro J, Sieiro C, Revilla G, González-Villa T, Zarra I (2001) Cloning and expression pattern of a gene encoding an α-xylosidae active against xyloglucan oligosaccharides from Arabidopsis. Plant Physiol **126:** 910–920

Schaffer R, Landgraf J, Accerbi M, Simon V, Larson M, Wisman E (2001) Microarray analysis of diurnal and circadian-regulated genes in Arabidopsis. Plant Cell **13:** 113–123

Segal E, Shapira M, Regev A, Pe'er D, Botstein D, Koller D, Friedman N (2003) Module networks: identifying regulatory modules and their condition-specific regulators from gene expression data. Nat Genet **34**: 166–176

Smith AM, Zeeman SC, Smith SM (2005) Starch degradation. Annu Rev Plant Biol **56:** 73–98

Smith AM, Zeeman SC, Thorneycroft D, Smith SM (2003) Starch mobilization in leaves. J Exp Bot **54:** 577–583

Smith SM, Fulton DC, Chia T, Thorneycroft D, Chapple A, Dunstan H, Hylton C, Zeeman SC, Smith AM (2004) Diurnal changes in the transcriptome encoding enzymes of starch metabolism provide evidence for both transcriptional and posttranscriptional regulation of starch metabolism in Arabidopsis leaves. Plant Physiol **136:** 2687–2699

Stessman D, Miller A, Spalding M, Rodermel S (2002) Regulation of photosynthesis during Arabidopsis leaf development in continuous light. Photosynth Res **72:** 27–37

Storey JD, Tibshirani R (2003) Statistical significance for genome-wide studies. Pro Natl Acad Sci USA **100:** 9440–9445

Stuart JM, Segal E, Koller D, Kim SK (2003) A gene-coexpression network for global discover of conserved genetic modules. Science **302:** 249–255

Tenorio G, Orea A, Romero JM, Merida A (2003) Oscillation of mRNA level and activity of granule-bound starch synthase I in Arabidopsis leaves during the day/night cycle. Plant Mol Biol **51:** 949–958

Thimm O, Blaesing O, Gibon Y, Nagel A, Meyer S, Krüger P, Selbig J, Müller LA, Rhee SY, Stitt M (2004) MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. Plant J **37:** 914–939

Tibshirani R, Walther G, Hastie T (2001) Estimating the number of clusters in a data set via the gap statistic. J R Statist Soc B **63:** 411–423

Williamson RE, Burn JE, Hocart CH (2002) Towards the mechanism of cellulose synthesis. Trends Plant Sci **7:** 461–467

Wingler A, Fritzius T, Wiemken A, Boller T, Aeschbacher RA (2000) Trehalose induces the ADP-glucose pyrophosphorylase gene, *ApL3*, and starch sythesis in Arabidopsis. Plant Physiol **124:** 105–114

Wurtele ES, Li J, Diao L, Zhang H, Foster CM, Fatland B, Julie Dickerson J, Brown A, Cox Z, Cook D, Lee E-K, Hofmann H (2003) MetNet: Software to build and model the biogenetic lattice of Arabidopsis. Comp Funct Genom **4:** 239–245

Yu TS, Zeeman SC, Thorneycroft D, Fulton DC, Dunstan H, Lue WL, Hegemann B, Tung SY, Umemoto T, Chapple A, et al (2005) α-Amylase is not required for breakdown of transitory starch in Arabidopsis leaves. J Biol Chem **280:** 9773–9779

Zeeman SC, Northrop F, Smith AM, ap Rees T (1998) A starch accumulation mutant of Arabidopsis thaliana deficient in a chloroplastic starch-hydrolysing enzyme. Plant J **15:** 357–365

Zeeman SC, Tiessen A, Pilling E, Kato KL, Donald AM, Smith AM (2002) Starch synthesis in Arabidopsis. Granule synthesis, composition, and structure. Plant Physiol **129:** 516–529

Zhou J, Loh YT, Bressan RA, Martin GB (1995) The tomato gene Pti1 encodes a serine/threonine kinase that is phosphorylated by Pto and is involved in the hypersensitive response. Cell **83:** 925–935

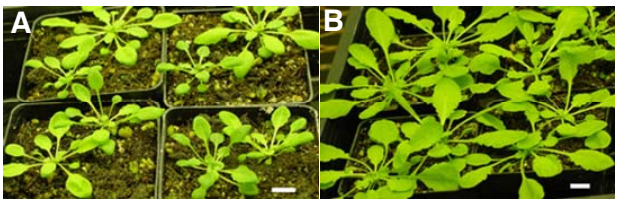


Figure 1. The overall phenotype of the antisense-*ACLA* and WT plants is quite similar, but leaves and rosette diameter of the antisense-*ACLA* plants are smaller. A, antisense-*ACLA* mutant. B, WT. Phenotype of antisense-*ACLA* and WT plants grown under SD conditions at 42 days after imbibition in the day time. Bar=1cm.

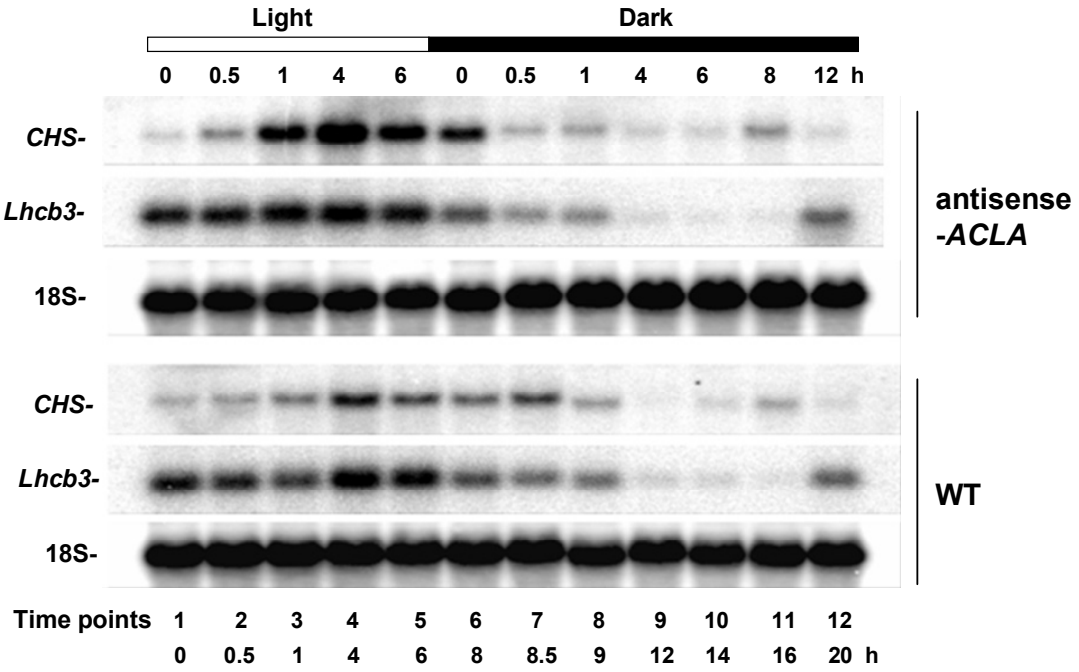


Figure 2. *Lhcb3* and *CHS* both display diurnal rhythms under SD and the circadian behavior of *Lhcb3* and *CHS* in the antisense-*ACLA* mutant was similar to WT. Northern blot analyses for *Lhcb3* and *CHS* transcripts in the antisense-*ACLA* and WT during SD diurnal cycle for replicate 1.

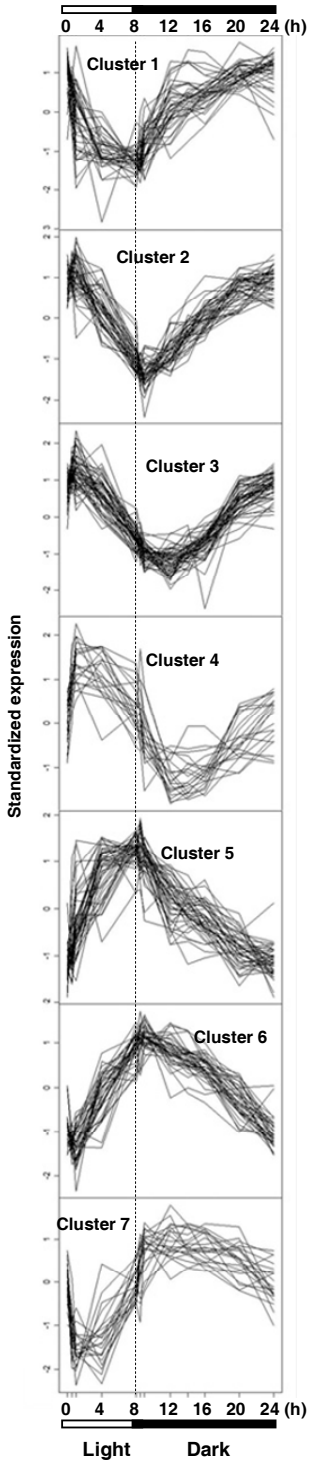


Figure 3. Clustering analysis of expression profiles of the genes that have the most significant transcript changes in WT across the SD diurnal cycle. Two hundred and seventy genes are significant when the FDR is controlled at the 0.00031 level. The 270 genes are grouped into seven clusters based on their expression patterns. Gene information is in Table S1 in Appendix E.

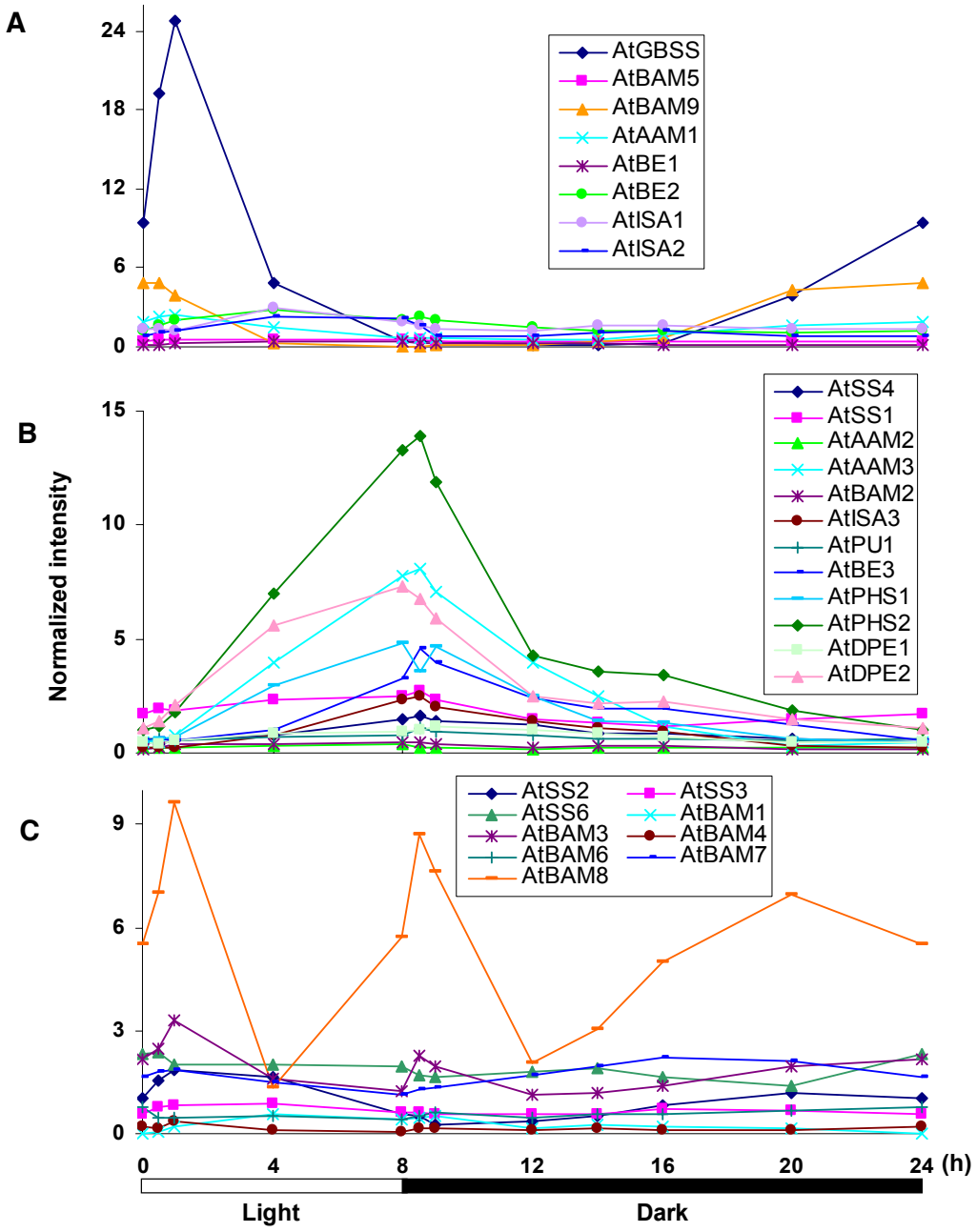


Figure 4. mRNAs of genes encoding enzymes involved in starch metabolism accumulate in three overall patterns: peak in the light; peak in the dark; peak in both the light and the dark. Overall expression profiles of starch metabolic genes in WT over SD diurnal cycle. The average normalized intensity for each chip is 1.

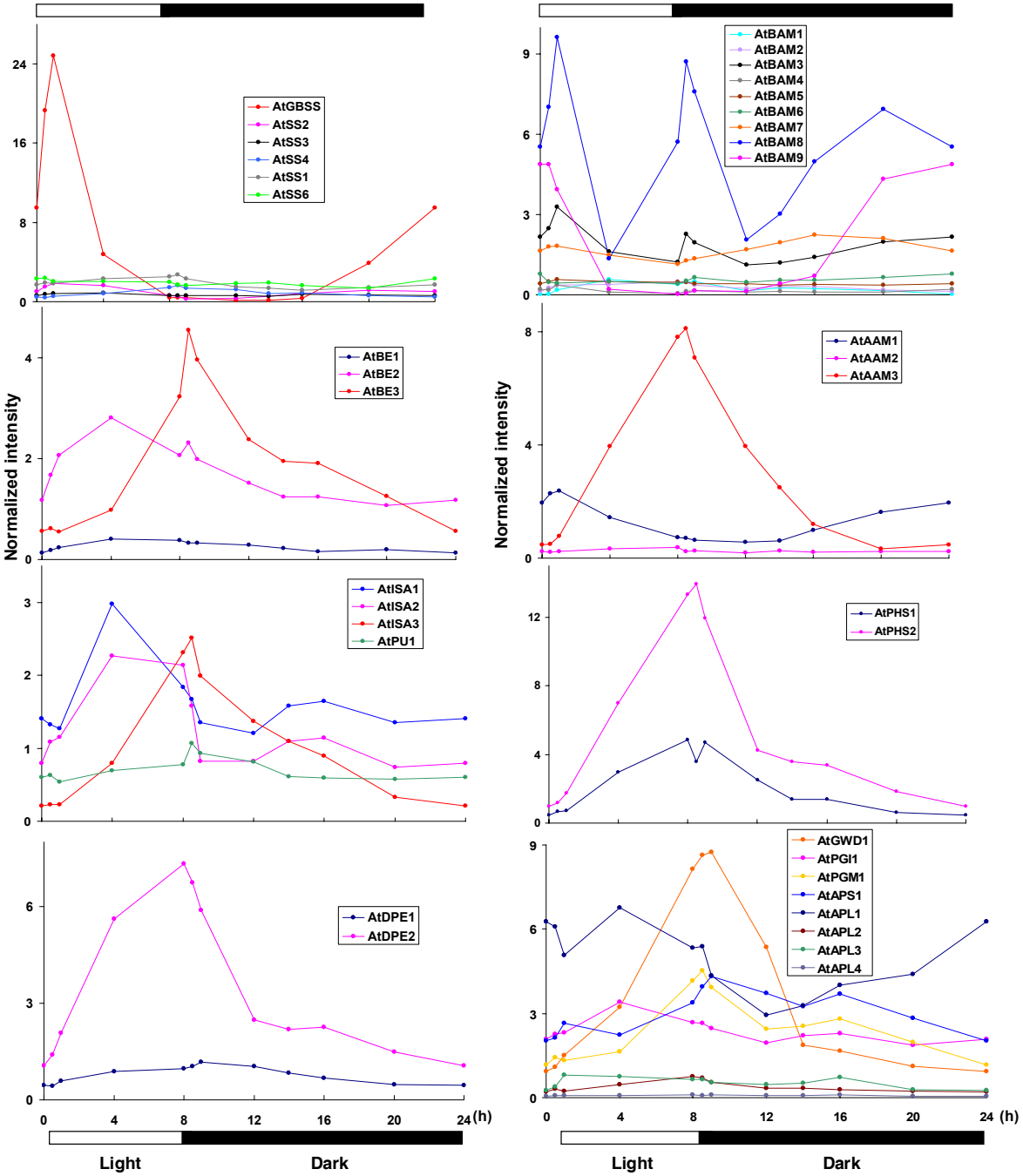


Figure 5. Expression profiles of genes encoding enzymes related to starch metabolism in WT over SD diurnal cycle. Genes are grouped by their biochemical activity. The average normalized intensity for each chip is 1.

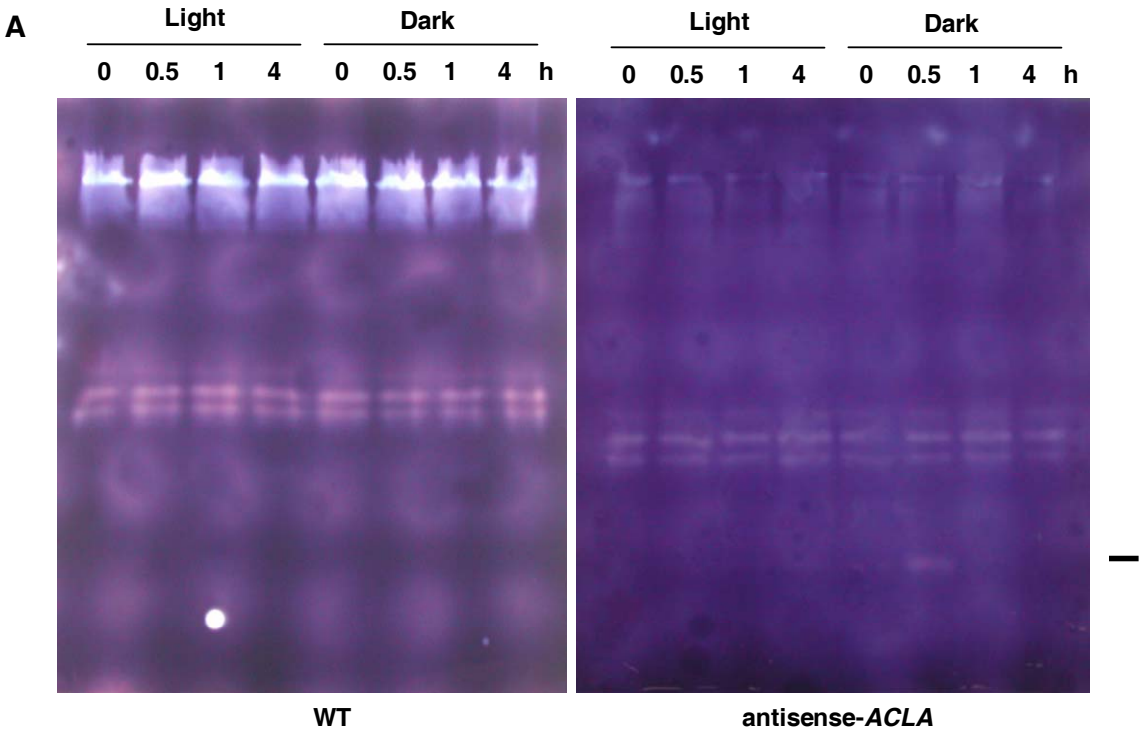


Figure 6. Native PAGE/activity gels (zymograms) of starch modifying and synthesizing activities in leaves of WT and antisense-*ACLA*. The enzyme activity does not reflect the fluctuations observed in transcript accumulation in response to the diurnal cycle. A, Total protein was separated on a native 7% polyacrylamide gel. Proteins were electroblotted to a starch-containing gel. Starch modifying enzyme activities were visualized by staining with I_2/KI. Activity of DBEs (AtISA2 and AtPU1- light blue), BEs (AtBE1, AtBE2, and AtBE3- red/orange). In antisense-*ACLA*, amylase activity is discerned at time points 4 h in light, 0 and 0.5 h in dark as a colorless band (-). B, Total protein was separated on a native 7% polyacrylamide gel containing 0.3% glycogen, incubated with ADP-glucose and stained with I_2/KI. Activity patterns of soluble starch synthases in WT and antisense-*ACLA* are similar and not noticeably changed by the diurnal cycle. The SS isoforms (AtSS3, AtSS1) are resolved from both WT and antisense-*ACLA* leaves. Two colorless bands (=) are detected in antisense-*ACLA* (4 h in light, 0, 0.5, 1, and 4 h in dark), but not in WT. Both experiments were repeated three times.

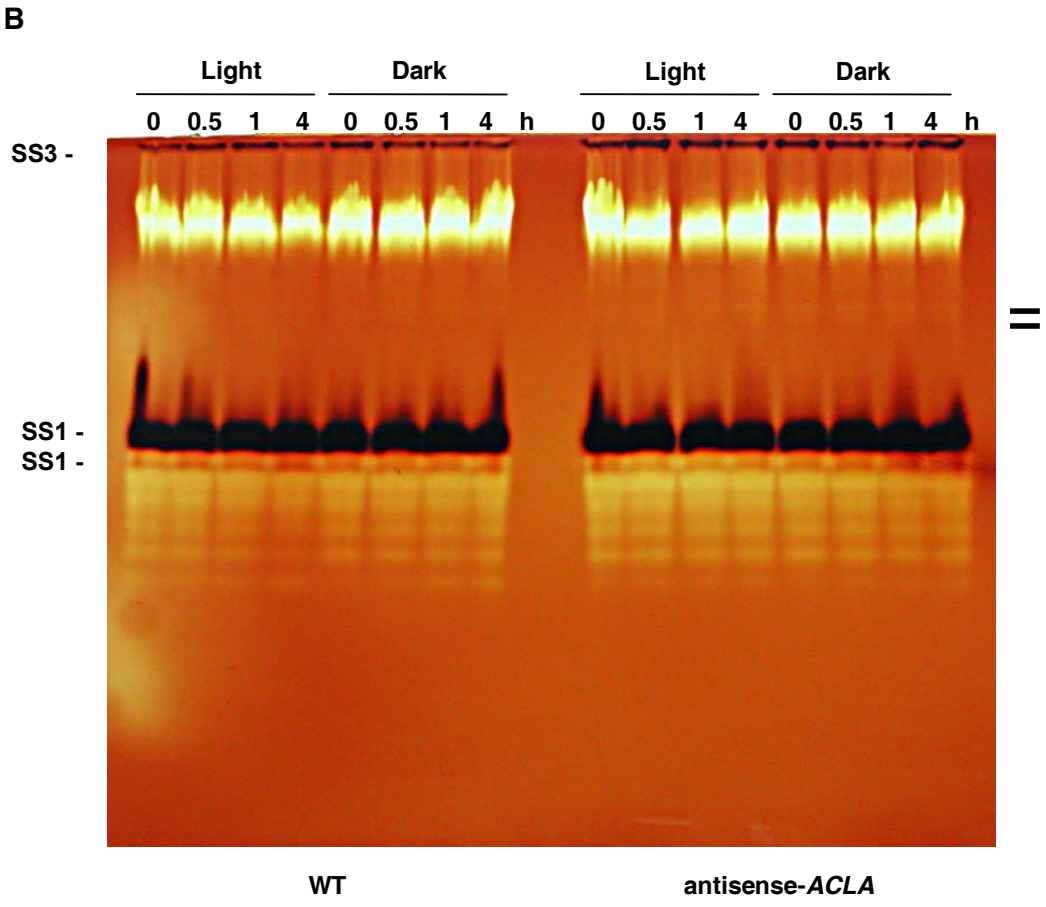


Figure 6. (continued).

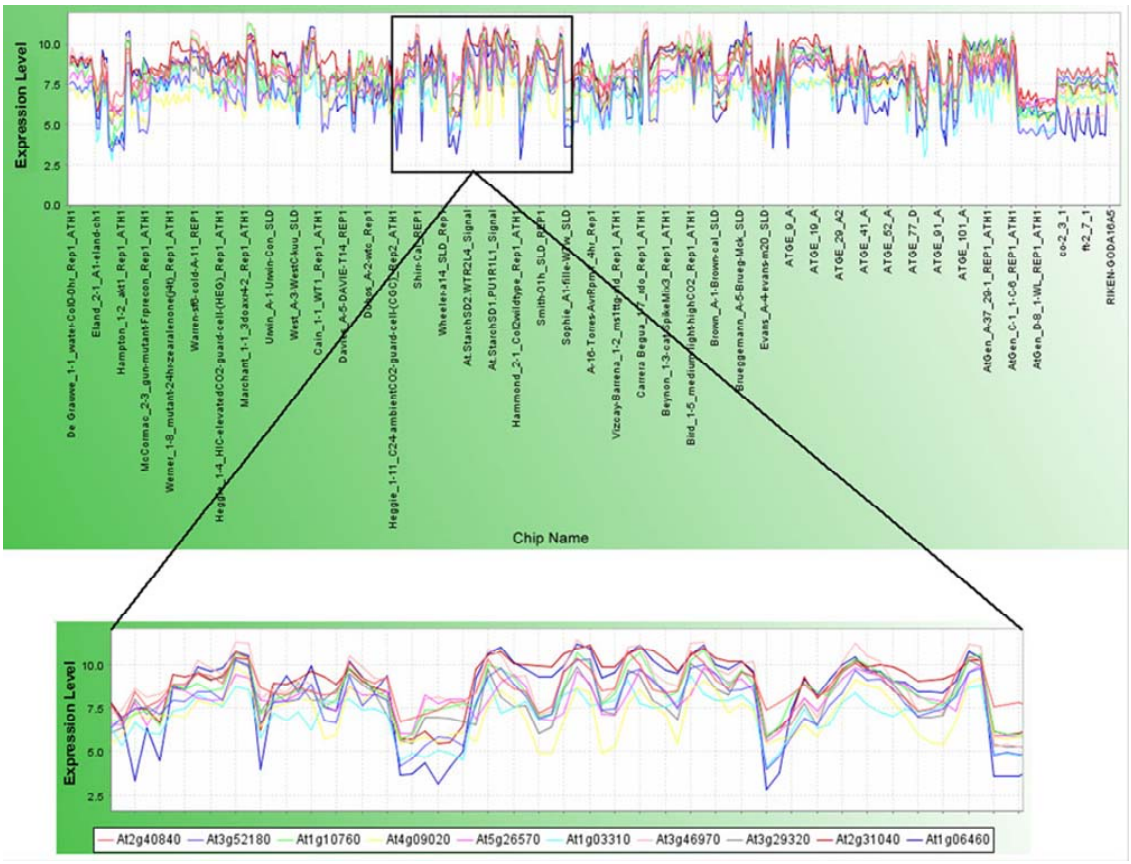


Figure 7. RNA accumulation of nine genes is similar to *DPE2*. Expression profiles across 963 chips for *DPE2*; other nine genes with the most similar RNA accumulation (*PTPKIS1*/*SEX4*, *GWD1*, *ISA3*, *GWD3*, *ISA2*, *PHS2*, *PHS1*, AT1G06460 and AT2G31040; correlation coefficients of 0.72 to 0.84), in MetaOmGraph. The average expression level for each chip is 6.6.

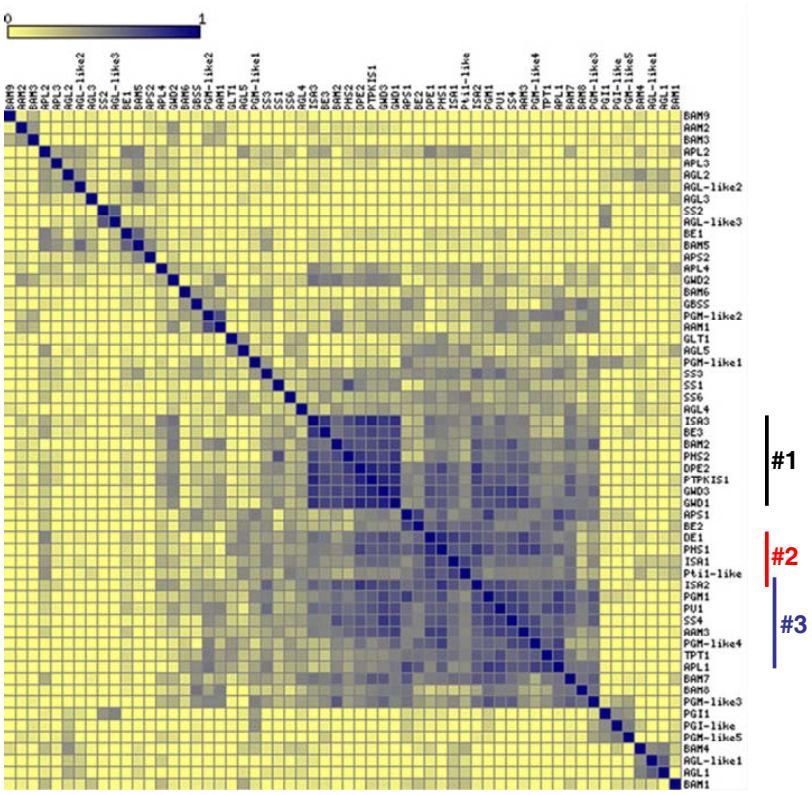


Figure 8. There are three groups of genes that are highly correlated in the starch metabolic genes. Correlations between starch metabolic genes. The values of Pearson correlations are color-coded. Correlation matrix was grouped by rows and columns to group genes with the highest correlation to each other together (darker squares along the diagonal). The vertical bars denote the groups.

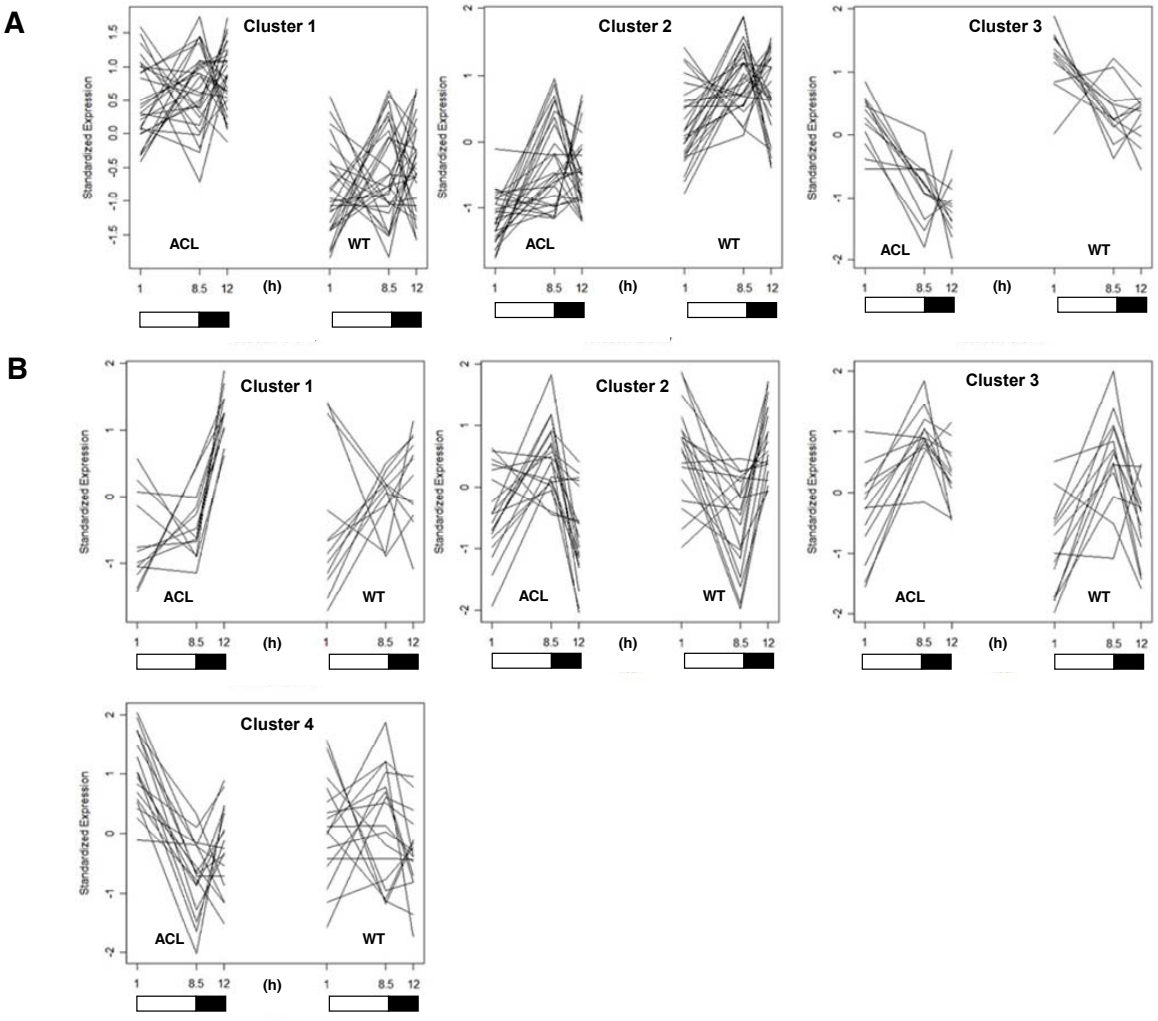


Figure 9 . Clustering analysis of expression profiles of the genes that have the most significant transcript changes between WT and antisense-*ACLA*. A, For the "Genotype" comparison between the WT and antisense-*ACLA;* gene information is in Table S2 in Appendix E. B, For the "Genotype*Time" comparison between the WT and antisense-*ACLA;* gene information is in Table S3 in Appendix E. Genotype "ACL" in this figure represents antisense-*ACLA*.

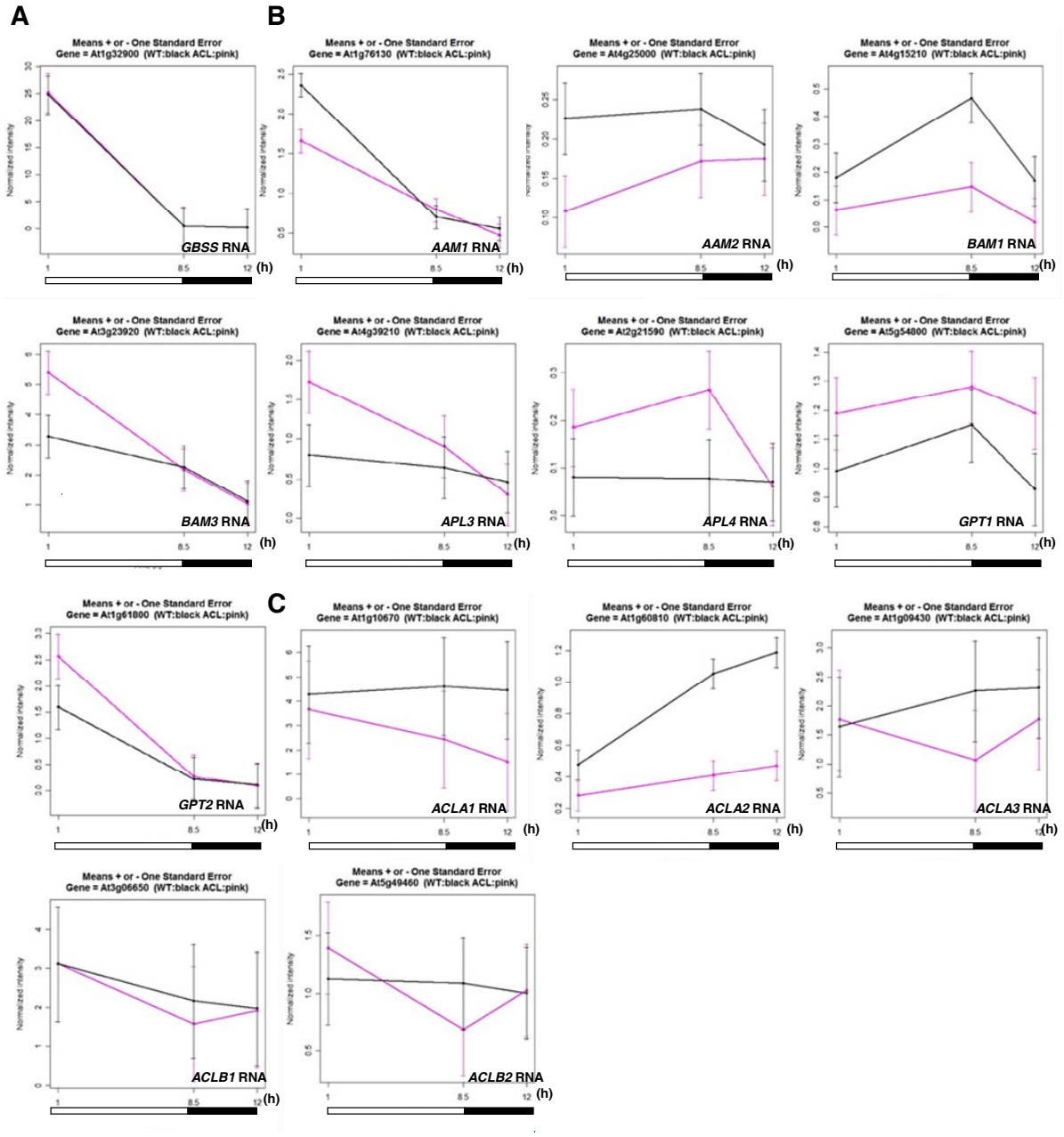


Figure 10. Expression profiles of genes involved in starch metabolism and encoding ACL in WT and antisense-*ACLA*. A, *GBSS* mRNA accumulation is almost identical in WT and antisense-*ACLA*. B, *AAM1*/*2*, *BAM1*/*3*, *APL3*/*4*, and *GPT1*/*2* RNA accumulations are different between WT and antisense-*ACLA*. C, Transcript accumulation of ACL genes in WT and antisense-*ACLA*; only *ACLA2* transcript is reduced in antisense-*ACLA*. The average normalized intensity for each chip is 1.

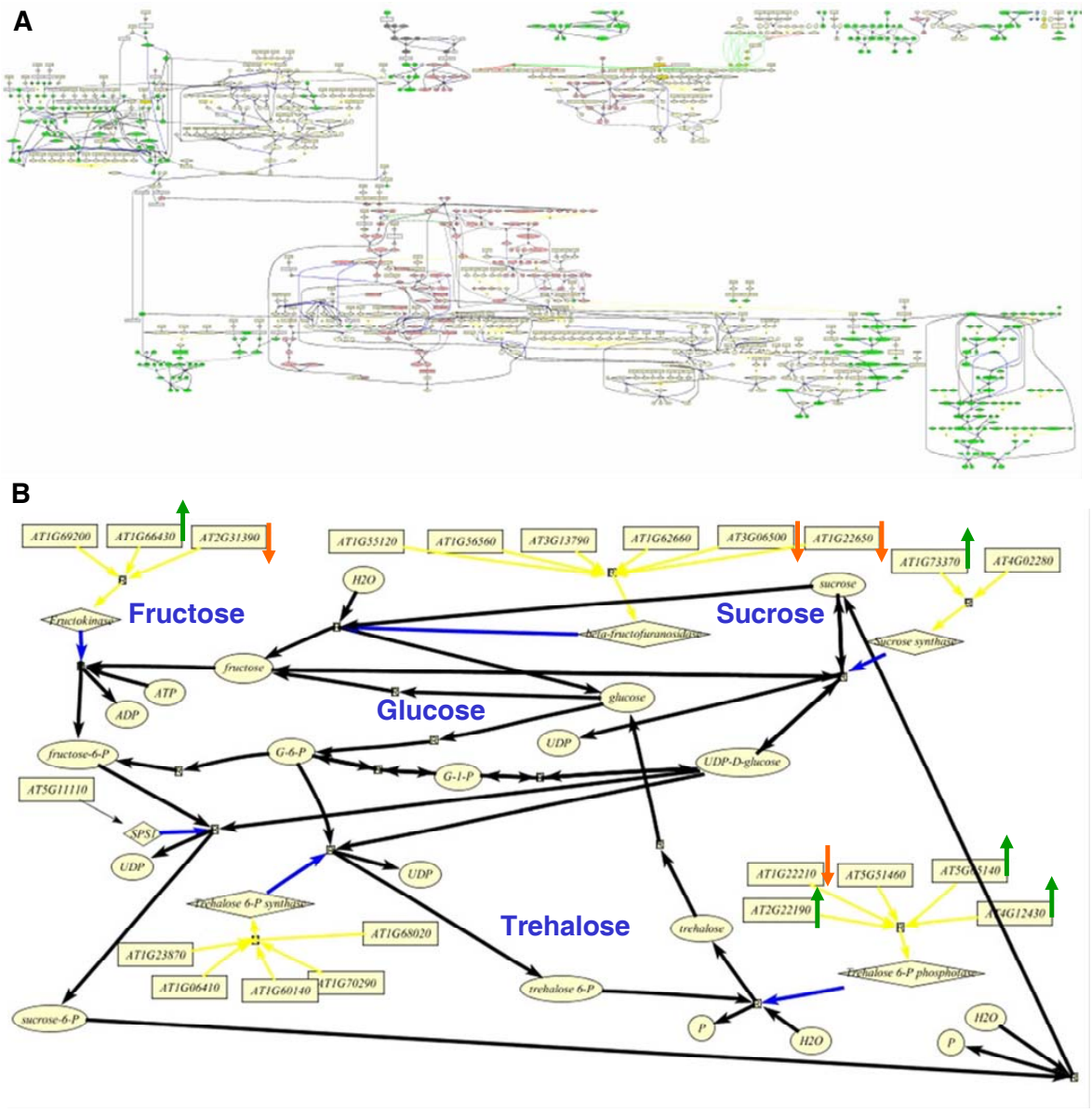


Figure 11. Metabolic networks from FCModeler. A, Overall network of carbohydrate and cell wall metabolism. B, Selected pathway to show the genes with altered RNA accumulation of sucrose and trehalose metabolism. Green arrow means this gene is up-regulated in antisense-*ACLA*, orange arrow means this gene is down-regulated in antisense-*ACLA*.

Table 1. Genes that are highly correlated to *ISA1* in RNA accumulation in MetaOmGraph.

Gene	Correlation coefficient	Locus ID
ISA1	1.00	At2g39930
ISA2	0.74	At1g03310
Pti1	0.73	At1g26150
DPE1	0.72	At5g64860
expressed protein	0.72	At4g10470
RCP1/*MEX1*	0.68	At5g17520
desulfhydrase	0.67	At1g48420
expressed protein	0.67	At3g60810
PHS1	0.66	At3g29320
CbbY like protein	0.66	At5g45170

Table 2. Pathways that have genes up-regulated or down-regulated in the antisense-*ACLA* mutant. The number of genes that have altered transcripts in that pathway is compared to number of genes total in the pathway, e.g., (6/21).

Pathway with genes up-regulated (41/173)	**Pathway with genes down-regulated (71/227)**
Cell wall metabolism (7/24)	***Cell wall metabolism (17/42)***
xylose degradation (1/3)	mannose degradation (1/2)
xylulose-monophosphate cycle (1/4)	GDP-D-rhamnose biosynthesis (5/16)
D-arabinose catabolism (1/4)	(deoxy)ribose phosphate metabolism (4/10)
dolichyl-diphosphooligosaccharide biosynthesis (2/10)	glucosinolate biosynthesis from homomethionine (2/5)
sinapate ester biosynthesis (2/3)	monoterpene biosynthesis (2/3)
Fatty acid elongation (12/32)	phenylpropanoid biosynthesis (3/6)
fatty acid elongation (3/5)	***Fatty acid oxidation (2/17)***
fatty acid omega-oxidation (1/2)	fatty acid beta-oxidation (2/17)
cutin biosynthesis (1/1)	***Sulfur containing amino acid biosynthesis (17/63)***
mannosyl-chito-dolichol biosynthesis (2/3)	sulfur assimilation (6/21)
mevalonate pathway (5/21)	homocysteine and cysteine interconversion (3/8)
Amino acid metabolism (8/33)	cysteine biosynthesis (2/13)
phenylalanine biosynthesis I (2/3)	methionine biosynthesis (3/12)
tyrosine biosynthesis II (2/4)	threonine biosynthesis from homoserine (3/9)
lysine biosynthesis (2/19)	***Hormone metabolism and signal transduction (27/74)***
serine biosynthesis (2/7)	jasmonic acid biosynthesis (5/18)
Hormone metabolism (2/8)	abscisic acid biosynthesis (4/6)
brassinosteroid biosynthesis (2/8)	cytokinins degradation (3/6)
Vitamins (3/11)	brassinosteroids signaling (4/12)
biotin biosynthesis (3/11)	APX1 signal transduction pathway (5/13)
Other (9/65)	lipoxygenase pathway (4/11)
chlorophyll Biosynthesis (3/17)	autophagy (2/8)
glyoxylate cycle (2/17)	***Other (8/31)***
glucose 1-phosphate metabolism (1/4)	glutamate degradation III (1/3)
de novo biosynthesis of pyrimidine ribonucleotides (3/27)	tryptophan biosynthesis (7/28)

Table 3. Fold change values of expression of genes involved in carbohydrate and cell wall metabolism for antisense-*ACLA* mutant. Fold change values were calculated by comparison to WT based on different metabolism pathways at three time points (1 h in the light (L1), 0.5 h (D0.5) and 4 h (D4) in the dark) respectively. Only genes with fold change values greater than 2 and error bars are not overlapped between genotypes are listed here. Expression in antisense-*ACLA* mutant is up-regulated (red fold change); expression in antisense-*ACLA* is down-regulated (blue fold change).

Genes up-regulated	Genes down-regulated
Starch synthesis ADPGlu-PP (*APL4*, At4g39210) **2** (L1) ADPGlu-PP (*APL3*, At2g21590) **3.4** (D0.5)	***Starch degradation*** *AAM2* (At4g25000) **-2** (L1) *BAM1* (At4g15210) **-3** (D0.5)
Trehalose trehalose-6-P phosphatase (At5g65140) **3** (L1) trehalose-6-P phosphatase (At4g12430) **2.1** (D0.5)	***Trehalose*** trehalose-6-P phosphatase (At1g22210) **-3** (L1, D4)
Cell wall synthesis UDP-galactose 4-epimerase (At2g34850) **6.2** (D0.5) cellulose synthase like gene (*AtCslA10*, At1g24070) **3.2** (D0.5) cellulose synthase like gene (*AtCslA03*, At1g23480) **2** (D4)	***Cell wall synthesis*** UDP-glucose/galactose 4-epimerase-like protein (At4g20460) **-2** (D0.5) mannose-6-phosphate isomerase 1 (At3g02570) **-2** (D0.5) rhamnosyltransferase, hemicellulose (At1g64910) **-2** (D0.5)
Cell wall modification beta-expansin (*EXPB1*, At2g20750) **2** (L1) beta-expansin(*EXPB3*, At4g28250) **3** (L1) **2.4** (D0.5) invertase/pectin methylesterase inhibitor (At2g47670) **3.7** (D0.5) expansin, putative (*EXP9*, At5g02260) **3** (D4) xyloglucan endotransglycosylase 1 (At3g23730) **5** (D4) xyloglucan endotransglycosylase (At4g03210) **5** (D4) xyloglucan endotransglycosylase (At1g10550) **4** (D4) endo-xyloglucan transferase (*XTR7*) (At4g14130) **3** (D4) pectin methylesterase (At5g47500) **2** (D4) invertase/pectin methylesterase inhibitor (At3g47380) **2** (D4)	***Cell wall modification*** xyloglucan endotransglycosylase (At5g65730) **-2** (L1) pectinesterase (At2g26440) **-3** (D0.5) pectinesterase (At4g33220) **-2** (D0.5) pectinesterase (At2g47550) **-2** (D4) pectinacetylesterase (At3g09410) **-2** (D4) invertase/pectin methylesterase inhibitor (At1g10770) **-3** (D4)
Cell wall degradation (1-4)-beta-mannan endohydrolase (At5g66460) **3** (L1) polygalacturonase (At1g10640) **3** (L1) pectate lyase (At1g67750) 2 (L1) **2.1** (D0.5) beta-xylosidase (At1g02640) **2** (D4)	***Cell wall degradation*** xyloglucan endo-1,4-beta-D-glucanase (*SEN4,* At4g30270) **-4** (L1) xyloglucan endo-1,4-beta-D glucanase (*XTR-3,* At4g30280) **-3** (L1) pectate lyase (At4g13210) **-2** (L1) polygalacturonase inhibiting protein 2 (*PGIP2,* At5g06870) **-2** (D0.5) xyloglucan endo-1,4-beta-D-glucanase (*XTR-6,* At4g25810) **-2** (D0.5) xyloglucan endo-1,4-beta-D-glucanase (At5g48070) **-2** (D4)

CHAPTER 4. A NOVEL GENE HIGHLY EXPRESSED IN THE *ATSS3* MUTANT

A paper to be submitted to Plant Cell

Ling Li, Dan Nettleton, and Eve Syrkin Wurtele

ABSTRACT

SS3 is one isoform of starch synthase involved in starch synthesis. The effects of eliminating SS3 on mRNA accumulations of other Arabidopsis genes were examined by global mRNA profiling between the *Atss3* mutant and WT of Arabidopsis. The genes with up-regulated transcripts are involved in transporting starch synthesis precursor or starch degradation product, in starch biosynthesis, in starch degradation, and in sucrose metabolism. Their altered transcript accumulation suggested a possible explanation to the *Atss3* mutant phenotype. A gene (*QQ*) was found with most significant transcript change in the "Genotype" comparison between the *Atss3* mutant and WT. It is a small protein (59 aa) with no sequence similarity to any other proteins. *QQ*'s expression in the background of WT, *Atss2*, *Atss3*, *Atss2*/*Atss3* and *Atpu1* mutants confirmed *QQ*'s high transcript accumulation in the *Atss3* mutant in microarray analysis. Iodine staining of the leaf starch of QQ RNAi seedlings indicates a starch excess at the end of light (16 h), and little difference at the end of dark (8 h) from WT. It is suggested that this unique small protein might be involved in regulating starch metabolism.

INTRODUCTION

Starch is a storage form of carbohydrate. It is very important for human diet and industry. Starch accumulates in leaf chloroplasts during the light time of the diurnal cycle. In the dark, leaf starch is converted into glucose and then made into sucrose, transported to other tissues for energy production or biosynthetic conversions. In Arabidopsis, starch granules accumulate transiently in developing seeds and then disappear. Mechanisms are not yet clear for starch biosynthesis, or its degradation to soluble glucose monomers. Starch granule formation and degradation are highly regulated processes. Amylose is composed of

α-(1→4) glucan chains. Amylopectin is composed of highly branched chain of α-(1→4) and α-(1→6) glycosidic bonds that are complex ordered. The ratios of amylose/amylopectin, and the branch structure of amylopectin are two major determinants of starch structure and characteristics. After the production of the glucosyl unit donor ADP-glucose, starch synthases (SS, encoded by 6 genes), branching enzymes (BE, encoded by 3 genes), debranching enzymes (DBE, encoded by 4 genes), α-amylases (AAM, encoded by 3 genes), β-amylases (BAM, encoded by 9 genes), disproportionating enzymes (DPE, encoded by 2 genes) and starch phosphorylases (PHS, encoded by 2 genes) are involved in the formation and degradation of starch. The loci for these genes are identified on the Arabidopsis starch metabolism network web site (http://www.starchmetnet.org). The significance of each gene family member is not yet understood.

SS contribute to starch formation by addition of glucosyl units to growing amylose or amylopectin chains through new α-(1→4) linkages. Synthesis of amylose from ADP-glucose may require SS only, whereas SS, BE, DBE and DPE may be involved in the synthesis of amylopectin (Myers et al., 2000). The six isoforms of starch synthase can be divided into 2 groups: granule bound starch synthase (GBSS) and soluble starch synthase (SSS). GBSS is responsible for amylose synthesis (Ball et al., 1998). At*ss1* loss-of-function mutant has altered amylopectin structure; SS1 may take part in synthesizing the amylopectin on the outer portion of the starch granules (Delvallé et al., 2005).

Atss3 loss-of-function mutant has an increased rate of starch biosynthesis, altered starch structure, and increased total SS activity and an increase in protein level of a non-SS3 SS (Zhang et al., 2005). This increase in SS activity despite the absence of SS3, indicated a possible negative regulatory function of SS3. Interestingly, SS3 may be regulated by the 14-3-3 proteins (Sehnke et al., 2001), a family of proteins that take important regulatory role in cellular signaling pathways (Sehnke et al., 2000).

Starch synthesis is part of a global network controlling carbohydrate flux in the plant. Sucrose can regulate starch metabolism. Sucrose is the main transported form of carbohydrate and starch is the major storage form in most plants. Sucrose-6-phosphate phosphohydrolase (SPP) catalyzes the synthesis of sucrose and the reaction is irreversible, stimulating sucrose synthesis (Lunn and REEs, 1990). β-fructofuranosidase (invertase, INV)

catalyzes the reaction to produce glucose and fructose from sucrose (Tang et al., 1996). Hexokinase (HXK) generates glucose-6-P from glucose. AtHXK1 is a glucose sensor and takes an important role in sugar-mediated regulation of gene expression (Moore et al., 2003). UDP-glucose pyrophosphorylase (UGPase) can convert UDP-glucose (which could be produced from sucrose) to glucose-1-P to provide carbon skeletons for biosynthesis of starch (Kleczkowski, 1996). ADP-glucose pyrophosphorylase (AGPase) generates ADP-glucose, the first step in starch synthesis. The regulatory subunit of AGPase is encoded by a family of four genes: *APL1*, *APL2*, *APL3* and *APL4* (Crevillén et al., 2003). Applied sucrose could induce *APL3* and *APL4* transcription in leaves (Crevillén et al., 2005).

To better understand function of SS3 and its potential role in starch metabolism and in regulation of starch metabolic network, leaves of the *Atss3* single gene knockout (Zhang et al., 2005) grown under short day conditions were used to generate a global mRNA accumulation profiles. Several genes related to starch biosynthesis had increased transcripts accumulation. We also found one gene (*QQ*) with most significantly altered RNA accumulation profile between the *Atss3* mutant and WT. We studied the gene's expression, trying to understand its function involved in starch metabolism.

RESULTS

Gene Expression in the *Atss3* Mutant and WT

The *Atss3* mutant has an increased rate of starch synthesis and altered starch structure (Zhang et al., 2005). To gain a better understanding of how loss of SS3 function affects expression of other genes of starch metabolism and other cellular processes, WT and *Atss3* mutant plants were grown under the short day (SD, 8 h light/16 h dark) diurnal cycle. Plants were grown and harvested according to a randomized complete block design with two independent biological replicates. Leaf RNA was analyzed using Affymetrix Arabidopsis ATH1 arrays. Microarray data were generated from Arabidopsis leaf tissue at five time points (1 h and 4 h in the light; half an hour, 4 h and 8 h in the dark). Initial profiling of starch metabolic genes under SD conditions shows most starch metabolic genes have peaks of transcription at these five time points over the diurnal cycle. The data were normalized by log (signal+1).

Gene expression was compared over the 5 time points between the *Atss3* mutant and WT. Seven hundred and sixty genes showed significant changes at the level of 0.01 based on the p-value. One hundred seventy five genes showed significant changes in expression between WT and Atss3 mutant in the genotype comparison when the false discovery rate (FDR) was controlled at the level of 0.162 using the method of Storey and Tibshirani (2003). Five clusters each contains genes with similar mRNA patterns (Figure 1) were generated for these 175 genes by K-medoids clustering method (see Materials and Methods). Genes in Cluster 1 (37 genes), and 4 (33 genes), are up-regulated in the *Atss3* mutant and genes in Cluster 2 (40 genes), and 3 (37 genes), are down-regulated in the *Atss3* mutant. Genes in Cluster 5 (28 genes) have peaks closer to the beginning of dark in the *Atss3* mutant.

The genes in each cluster, together with p-value, q-value, TAIR gene annotation (http://www.arabidopsis.org) and MapMan category (Thimm et al., 2004) are listed in Table S1 (Appendix F). Twenty-30% of genes in each cluster are annotated as "expressed protein" with unknown function. Genes involved in carbohydrate metabolism, regulation of transcription, and protein metabolism (including posttranslational modification and degradation) are dispersed among each cluster. The Cluster 1, 2, 4 and 5 all contain genes involved in cell wall metabolism. Additionally, Cluster 1 and 2 contain genes involved in transport, hormone metabolism and stress response; Cluster 3 contains several genes involved in hormone metabolism, stress response and signaling; Cluster 4 contains genes involved in phosphate transport and signaling; Cluster 5 contains genes involved in hormone metabolism and stress response. A typical example of the expression pattern of a gene from each cluster is shown in Figure 2.

Gene Expression Related to Starch Metabolism in the *Atss3* Mutant and WT

Clusters 1, 3 and 5 contain genes known to be involved in starch metabolism: *GBSS*, Cluster 1; *BAM9* and *SS3,* Cluster 3; *ISA3*, Cluster 5. The *Atss3* mutant has been reported to have increased total SS activity and the rate of starch synthesis (Zhang et al., 2005). To determine the effect of loss of SS3 on starch metabolism, we checked the RNA accumulation of genes involved in or related to starch metabolism. If at any time point, the error bars of gene expression are not overlapped between the genotypes, this gene is considered as

expression altered at that time point. Twenty two genes proven or thought to be involved in starch and sucrose metabolism have altered transcripts in the *Atss3* mutant relative to WT. These genes can be divided into four groups: transportation of starch synthesis precursor (G-6-P) and starch degradation products (triose phosphate and glucose) across the membranes of the plastids; starch synthesis; starch degradation; and sucrose metabolism (Figure 3).

Of the genes encoding glucose/triose transporters, *GPT2* (glucose-6-phosphate/phosphate translocator) and *TPT1* (triose phosphate/phosphate translocator) are up-regulated in the *Atss3* mutant; while *GLT1* (glucose transporter) is down-regulated in the *Atss3* mutant (Figure 3A). Of the starch synthesis genes, six genes (*GBSS*, *BE3*, *ISA2*, *APL1*, *APL2*, and *APL3*) are up-regulated in the *Atss3* mutant relative to WT (Figure 3B); *SS6* is down-regulated in the *Atss3* mutant. Of the starch degradation genes, *BAM9* (unknown function, Smith et al., 2005) is down-regulated in the *Atss3* mutant, *ISA3* (Delatte et al., 2006), and *SEX4* (Niittylä et al., 2006) are up-regulated (Figure 3C).

Eight sucrose metabolic genes: *SPP2*, *SPP4* (sucrose phosphate phosphatase), *INV1* (invertase), *HXK1-4* (hexokinase), and *UGP* (UDP-glucose pyrophosphorylase) transcripts are up-regulated in the *Atss3* mutant (Figure 3D); At4g26530 (cytosolic fructose-bisphosphate aldolase, FBPALD) is down-regulated in the *Atss3* mutant (Figure 3D).

A Gene of Unknown Function is Highly Expressed in the *Atss3* Mutant

Of the genes that have the significant changes in transcript level in the "Genotype" comparison between the *Atss3* mutant and WT over the SD diurnal cycle, one particular gene attracted our interest. At3g30720 (we name it *QQ*), has the most significant transcript change. It is up-regulated in the *Atss3* mutant, showing more than a 7-fold change in accumulation for all 5 time points over the diurnal cycle (Figure 4). This gene also has the lowest p value ($8.59E^{-09}$) in the "Genotype" comparison between the *Atss3* mutant and WT. In the WT, accumulation of *QQ* mRNA decreases from 1 h in the light till it reaches its bottom at 4 h in the light. It increases slowly thereafter and increases more after the dark comes. It has a peak at 4 h in the dark and then decreases over the diurnal cycle. In the *Atss3* mutant, *QQ* mRNA accumulation decreases in the light, reaches its lowest expression level after half an hour in the dark, and increases throughout the dark period.

RT-PCR confirms the up-regulation of *QQ* in the *Atss3* mutant. *QQ* RNA accumulates more in the *Atss3* mutant than in WT (Figure 5). The reaction was repeated at two annealing temperatures: 47°C and 51°C.

QQ is annotated on TAIR as an "expressed protein", with "cellular component unknown, molecular function unknown, biological process unknown". *QQ* shows no nucleotide sequence similarity to any other gene of Arabidopsis or of any other organisms with sequenced genomes. The predicted protein is tiny, only 59 aa, and also does not show any sequence similarity to any other protein in Arabidopsis or in any other organism. The transcript starts with a 200-bp 5'UTR. The gene contains two introns, one in the 5'UTR (114 bp between upstream -29 and -30 from ATG codon), and one in the protein coding sequence (86 bp between position 18 and 19 from ATG codon) (Figure 6).

QQ contains five transcription binding motifs in the 754-bp intergenic region upstream of its 5'UTR (Table 1): one light responsive motif (CCA1 binding site motif), two protein-protein interaction motifs (TGA1 binding site motif and UPRMOTIFIAT), and two ABA responsive motifs (ABRE-like binding site motif and ACGTABREMOTIFA2OSEM). *QQ* contains a "CAACTG" MBS motif (MYB transcription factor family binding site involved in drought-inducibility) at -25 from ATG in the 5'UTR.

QQ is located amidst a cluster of small genes with unknown functions, annotated as "expressed protein" or "hypothetical protein", and transposon like genes (not as small as *QQ*) annotated as "gypsy-like retrotransposon family", or "CACTA-like transposase family" (Figure 7, Table 2).

To get an idea of *QQ*'s expression and to postulate its possible function, we checked a co-normalized set of public microarray data across 72 microarray experiments, representing a wide range of developmental conditions, and environmental and genetic perturbations (Mentzen, 2006). These data were analyzed and visualized using MetaOmGraph (MOG), a software developed to plot and analyze large sets of data (http://metnetdb.org/MetNet_MetaOmGraph.htm). Only three of the genes listed in Table 3 (*QQ*, At3g30730, and At3g30740) have probe sets on the Affymetrix ATH1 array. Of these, *QQ* is the only gene that is expressed at a level sufficient for detection (Figure 8A). *QQ* RNA is highly accumulated in cell cultures with auxin added (Figure 8B), and in mature pollen

(Figure 8C and D). Its RNA accumulation increases as pollen develops and reaches a peak when pollen matures. In addition, *QQ* expression is up-regulated (about 2-3 fold change) in cell suspensions where there is no sucrose in the culture.

Spatial Expression Patterns of *QQ*

To investigate the spatial localization of *QQ* to try to gain insight into its function, we constructed a vector that expressed fused GFP/GUS under the control of the *QQ* promoter region (Figure 6). The promoter::GFP/GUS construct was transformed into Arabidopsis (WT), the *Atss2*, *Atss3*, *Atss2*/*Atss3*, and *Atpu1* mutants. We included the 3 soluble SS mutants since *QQ* RNA accumulation is altered in the *Atss3* mutant. We wanted to compare *QQ* expression in these different genotypes to see the effect of gene loss on *QQ* expression. We also used the *Atpu1* mutant as a negative control. At least ten independent transgenic lines for WT in the T2 generation were screened. The lines were assessed at various stages of development from germination to senescence.

In WT, *QQ* is expressed in young seedlings, hypocotyl, cotyledons mesophyll cells and vasculature. It is also expressed in leaf mesophyll cells and vasculature, hydathodes, trichomes, petioles, and in the shoot meristem. It is highly expressed in root: the vasculature, root tip and throughout the vasculature and cortex of the old roots. It is also expressed in emerging lateral root. It is expressed in flower sepals, filaments, pollens, stigma, style, ovary wall, receptacle, and in the ovule and young silique wall (Appendix F, Figure S1). The pattern of expression of *QQ* in the *Atpu1* mutant is indistinguishable from that of WT (Appendix F, Figure S2).

QQ has differing expression among the starch synthasis mutants (Appendix F, Figure S2). In the *Atss3* mutant, *QQ* RNA accumulation is higher than in WT. The *Atss2* mutant has lower *QQ* RNA accumulation than WT (or the *Atss3* mutant). *QQ* RNA accumulation pattern in the *Atss2*/*Atss3* double mutant is similar, but slightly reduced, compared to the *Atss3* mutant. In the *Atss3* and *Atss2*/*Atss3* mutants, *QQ* is expressed in petal and funiculus (Figure 9 and Appendix F, Figure S1, S2). Appendix F, Table S2 includes the summary of *QQ* expression. Figure 9 gives a comparison of *QQ* expression in WT, the *Atss3*, *Atss2*/*Atss3*, *Atss2* mutants in leaves, flowers, pollen, and silique.

Sub-cellular Localization of *QQ*

QQ is such a unique small protein. To better understand QQ's function, we made construct to express GFP/GUS fused to the N-terminal region of QQ, under the control of the *QQ* gene promoter region. Since QQ is very small with 59 aa, we included 259 bp (including the 86-bp intron between position 18 and 19, 57 aa) (Figure 6).

PSORT (http://psort.hgc.jp/) predicts QQ to be in the cytoplasm with certainty=0.450. In leaf, QQ protein appears not to be in the plastids or the nucleus (Figure 10).

Reduced *QQ* Transcript Affects Starch Metabolism

Although *QQ* is increased in the Atss3 and Atss2/Atss3, both mutants with increased starch (Zhang, 2006), it is unclear what role if any it might play in starch accumulation. To study whether the loss of transcript of *QQ* will affect the plant in particular starch metabolism, we generated a QQ RNAi mutant. We made a construct that inserted piece of *QQ* which is 150 bp from position 24 to 173 into pB7GWIWG2(II) (Figure 6).

Under a 16 h light/8 h dark LD conditions, QQ RNAi mutants grow faster than WT (Figure 11A). Since *QQ* has high transcript in the pollen, we suspected at first that QQ might affect male transmission. To determine whether this was the case, we checked the survival rate of QQ RNAi mutant. For one QQ RNAi line, 20 T2 seedlings died and 62 T2 seedlings survived after herbicide selection. Chi-square test was performed for the fitness test for the segregation ratio. The p-value of the chi-square test is 0.897791. It indicates that, though *QQ* is expressed in pollen, there is nothing wrong with male transmission in the QQ RNAi mutant. RT-PCR confirms that the transcript of *QQ* is reduced in the QQ RNAi mutant (data not shown). The QQ RNAi mutant grows faster than WT under the LD conditions (Figure 11A). Iodine staining of the leaf starch of seedlings indicates a starch excess at the end of light (16 h) (Figure 11B), and little difference at the end of dark (8 h) (data not shown) in the QQ RNAi mutant compared to WT. This result indicates that reduced *QQ* transcript does affect starch metabolism, rather surprisingly resulting in more starch content in QQ RNAi mutant at the end of light.

DISCUSSION

There are several families of genes important for starch metabolism, for example, starch synthases (SS), debranching enzymes (DBE), branching enzymes (BE), β-amylases (BAM), α-amylases (AAM), phosphorylases (PHS), and disproportionating enzymes (DPE). SS, BE, DBE and DPE may be involved in the synthesis of starch (Myers et al., 2000). In the *Atss3* mutants, loss of SS3 may cause changes that affect the transcription of certain genes that are sensitive to those changes. These changes might provide possible clues as to the mechanism in bringing about the starch excess phenotype. The *Atss3* mutant has increased starch biosynthesis rate and total SS activity, altered leaf starch structure, and an up-regulated specific SS activity (Zhang et al., 2005). Otherwise, the visible phenotype resembles the WT. In the *Atss3* mutant, multiple genes are involved in transporting starch synthesis precursor or starch degradation product, starch biosynthesis, starch degradation, and sucrose metabolism, have altered expression (Figure 12).

Glucose 6-phosphate/phosphate translocator was proposed to import glucose-6-P into plastids to be used as a precursor for starch or to take part in other metabolism in plastids. *GPT1* and *GPT2* (At1g61800) could rescue *pgi1* mutant from low starch in leaf (Niewiadomski et al., 2005). The increased transcript of *GPT2* in the *Atss3* mutant might suggest a possible starch synthesis induction by increased *GPT2* RNA to import more glucose-6-P to be used for starch synthesis in plastids. The TPT, which is located in the chloroplast inner envelope membrane, exchanges chloroplastidic triose-phosphates (triose-P, generated from CO_2 fixation) with cytosolic orthophosphate (Pi) during photosynthesis. This supplies the cytosol with the precursors for sucrose synthesis (Walters et al., 2004). The *tpt-1* mutant has increased starch in the light, and induced starch mobilization in the dark (Schneider et al., 2002). The *ape2* mutant loss of TPT function has increased starch and *GPT2* transcript is decreased in the *ape2* mutant (Walters et al., 2004). The significance of the up-regulated *TPT1* transcript in the *Atss3* mutant is not clear yet. Trethewey and ap Rees (1994) reported a mutant which is unable to transport glucose across the chloroplast membranes contains more starch at the end of light than WT. The down-regulated of *GLT1* in the *Atss3* mutant which contains increased starch is consistent with their report.

Trehalose and sucrose have been hypothesized to induce starch synthesis in Arabidopsis source tissue via an increase in ADP-glucose pyrophosphorylase activity and *APL3* expression (Wingler et al., 2000; Rook et al., 2001). Kolbe et al. (2005) reported external feeding trehalose could induce starch synthesis; trehalose phosphate synthase increased leaf starch. A trehalose-6-phosphate synthase (*TPS*, At4g17770) and two trehalose-6-phosphate phosphatase (*TPP*, At5g51460 and At4g12430) have increased RNA in the *Atss3* mutant (data not shown). In addition, the *APL3* transcript was up-regulated, in a similar pattern to that of *GBSS*. The *APL3* mRNA is increased in at 4 h in the light and half an hour in the dark. It is also possible that loss of SS3 increased *TPS* RNA accumulation and affected the sucrose or trehalose content. The sucrose or trehalose induced starch synthesis, accompanied with increased of *APL3* RNA accumulation and increased ADP-glucose pyrophosphorylase activity. Of the starch biosynthetic genes with altered transcripts in the *Atss3* mutant, the increased *GBSS* transcript might increase the starch synthesis and the total SS activity. The specific up-regulated SS activity described in Zhang et al. (2005) could be that of GBSS. The significance of *SS6* down-regulation is unknown. BE3 is a redundant isoforms of BE2, the elimination of BE2 and BE3 functions results in starch decrease in Arabidopsis (Dumez et al., 2006). The up-regulated *BE3* in the *Atss3* mutant is consistent with its involvement in starch synthesis. ISA1 and ISA2 are suggested to form a protein complex with ISA2 as the regulatory subunit, involved in starch synthesis (Hussain et al., 2003; Delatte et al., 2005). Though the transcript of *ISA1* is not altered, up-regulated *ISA2* might regulate the activity of the ISA1/ISA2 protein complex and contribute to the increased starch synthesis and altered amylopectin structure in the *Atss3* mutant.

The genes we previously discussed are involved in or related to starch synthesis. Interestingly, *ISA3*, a gene suggested to be involved in starch degradation (Delatte et al., 2006), has increased transcript in the *Atss3* mutant, together with *ISA2*, its possible co-functioning gene (Chapter 2 in this dissertation). The increased *ISA2* and *ISA3* transcripts might indicate that they take part in altering the amylopectin structure, rather than degradation.

Sucrose metabolism is tightly related to starch metabolism. Quite a few genes involved in sucrose metabolism have altered transcripts in the *Atss3* mutant. The up-

regulated *SPP2*/*4*, *INV1*, *HXK1-4* and *UGP* transcripts in the *Atss3* mutant reflect a direction of change toward increase of G-6-P, which could be transported to chloroplast as a precursor for starch biosynthesis at the transcriptional level. The significance of the down-regulated fructose-bisphosphate aldolase transcript remains to be clarified.

In the *Atss3* mutant, loss of SS3 results in increased transcript of several genes involved starch synthesis (for example, *GBSS*, *APL2*, *APL3*). It is consistent with the possible negative role of SS3 on starch synthesis as proposed by Zhang et al. (2005). The increased *GBSS*, *APL3* and increased starch synthesis present a partial explanation for the postulated negative role of SS3 in starch synthesis.

Several genes that are significantly differently expressed between the *Atss3* mutant and WT are involved in cell wall metabolism (Figure 2 and Appendix F, Table S1). This might indicate the carbon partitioning into starch causes changes in cell wall metabolism. Also, the *Atss3* mutant grows a little bit slower and looks smaller than WT under SD (data not shown). Whether there is relation between the changed expression of genes involved in cell wall metabolism and the smaller size of *Atss3* mutant under SD remains to be tested.

QQ is a tiny protein of only 59 aa. It is unique with no sequence similarity to any other protein in Arabidopsis or in any other organism. Several *QQ* nearby genes are also very short with unknown function. It is possible that these genes might have similar function or take part in similar biological process. Five of the 14 genes with known function that are correlated with *QQ* (50% and above, in MetaOmGraph) are involved in RNA recognition or are transcription factors (Table 4). It is possible that either *QQ* might be under the control of these transcription factors or QQ might have similar functions with these transcription factors.

The epsilon 14-3-3 protein is implicated to function together with phosphorylated SS3 in regulation of starch accumulation and reduced epsilon 14-3-3 protein in the antisense mutant results in increased starch (Sehnke et al., 2001). *QQ* is highly expressed in the *Atss3* mutant; we initially suspected that *QQ* belongs to 14-3-3 protein family. However, there is no protein sequence similarity between *QQ* and 14-3-3 proteins. Also, *QQ* is not highly correlated to 14-3-3 protein genes in RNA accumulation profiles. It is not likely that QQ is a 14-3-3 protein. But the transcript of epsilon 14-3-3 protein is decreased at 1 h in the light in

the *Atss3* mutant. Of the 13 genes of the Arabidopsis 14-3-3 proteins, five are up-regulated, three are down-regulated, four have transcripts not altered, and one is not represented on the ATH1 chip. The loss function of SS3 in the *Atss3* mutant also affects the transcript level of 14-3-3 proteins. The role of the change in expression of 14-3-3 genes, especially the down-regulation of epsilon 14-3-3 gene (1.3-fold change) in the up-regulation of QQ is still not clear.

The expression of *QQ* is not low in WT. The spatial expression pattern of *QQ* in leaves and pollen is consistent with what we found in microarray analysis. *QQ* is highly up-regulated in the *Atss3* mutant, in the *pen3* mutant (ATP binding cassette transporter, Stein et al., 2006), in *etr1* (ethylene receptor) loss-function mutant and *nahG* (salicylate hydroxylase converts salicylic acid (SA) to catechol) over-expression mutant, in mutants with over expressed gene involved in cytokinin-degradation and microRNA, in mutant loss gene function involved in RNA-binding (Genevestigator, https://www.genevestigator.ethz.ch/at/index.php). QQ transcript responds to high CO_2 (Stein et al., 2006). The reduced QQ transcript in the QQ RNAi mutant contains more starch relative to WT. This shows evidence that QQ is involved in starch metabolism. Since QQ contains no known protein motif, it is not likely that QQ has catalytic function. QQ might be involved in plant defense response, signaling, RNA binding or RNA regulation.

No paper has reported an involvement of such a small gene, like *QQ*, in starch metabolism. QQ is unique. To clarify its function in regulating starch metabolism will shed light on understanding starch metabolism.

MATERIALS AND METHODS

Microarray Experiment

Plants used in this study were wild-type *Arabidopsis thaliana* (ecotype Columbia), the *Atss3* mutant and the *pu1* mutant. The mutants are both in Columbia background. Seeds were surface sterilized in a 50% bleach/ 0.02% Triton X-100 solution for 7 min and rinsed in sterile distilled H_2O three times. The growth medium was buffered with 2.56 mM MES at pH 5.7 and contained 4.3 g/L Murashige and Skoog salts (GibcoBRL, Life Technologies, Rockville, MD), 1% sucrose, 1X B5 vitamins (100 μg/mL myo-inositol, 1 μg/mL pyridoxine

hydrochloride, 1 μg/mL nicotinic acid, 10 μg/mL thiamine hydrochloride), and 0.6% (w/v) agar gel (Sigma, St. Louis, MO). The seeds were placed in a growth chamber (66.6% RH at 24.2°C) under a short day (SD) photoperiod (8 h light/16 h dark) from fluorescent lamps (132.8 $\mu M \cdot m^{-2} \cdot s^{-1}$ PAR). After 21 d, seedlings were transplanted to a sterile potting medium (LC1 Sunshine Mix, Sun Gro, Horticulture, Inc., Bellevue, WA) and fertilized with 1X Nutriculture Grower's Special Blend 21-8-18^{PLUS} (Plant Marvel Labs, Chicago, Il). Plants were arranged in randomized blocks using R (www.r-project.org) and returned to the same growth chamber and environmental conditions as used for germinating seeds. It was Randomized Complete Block Design for 3 genotypes (WT/*Atpu1*/*Atss3*, we initially planted, harvested and processed plants of these three genotypes together, but we analyzed microarray data of *Atss3* and WT only in this paper). The plants were planted in a row (randomized in three pots) with 7 rows in each flat. There are 21 pots in one flat with 2 plants / pot. There are totally 112 rows in 16 flats. Randomized 8 rows were harvested for each time point from different flats and 12 time points (0, 0.5, 1, 4, 6, 8, 8.5, 9, 12, 14, 16, and 20 h) were harvested. We selected 5 time points (1, 4, 8.5, 12, and 16 h) for microarray data. Rosette leaves (#5 to #8) were harvested from six-week-old seedlings during the SD photoperiod under a green light. Each sample consisted of leaves from sixteen plants. Samples were frozen in liquid N_2 immediately after harvest and stored at -80°C for RNA extraction. The experiment was done twice. Different randomizations for plant growth and harvest were used for two replicates.

Microarray Analysis

The microarray data were normalized by log (signal+1). We fit a mixed linear model with fixed effects for time and genotype and random effects for replicate to identify genes that exhibit significant change in expression across the 5 time points per genotype and 2 genotypes surveyed. We used the method of Storey and Tibshirani's (2003) to estimate FDR to gain more power for detecting differences when comparing the WT and mutant genotypes. Thresholds for significance based on q-values were selected to maintain false discovery below desired rates. After matrix of estimated expression means was obtained, the matrix of means was standardized so that Euclidean distance would be related to 1-correlation distance.

K-medoids clustering was conducted with the cluster number that Gap analysis recommended. The Gap statistic method (Tibshirani et al., 2001) calculates the number of clusters by comparing the degree of clustering of the data to that of a simulated uniform (cluster-free) reference distribution.

Construction of *QQ* Promoter/Promoter+Target::GFP/GUS Fusion and *QQ* RNAi Vectors

QQ promoter/promoter+target::GFP/GUS fusion and RNAi constructs were made for *QQ* by cloning amplified promoter/promoter+target regions into a binary vector. The upstream intergenic region for *QQ* is 754 bp. A single construct was made for *QQ* promoter::GFP/GUS fusion: QQp (upstream 733 bp from -1). Primers used were QQptarF (5'-AAAATCTGCAATTATGTAAAC-3'), QQpR (5'-GGTGTTTGGTTCTTAGATTC-3'). A single construct was made for *QQ* promoter+target::GFP/GUS fusion: QQptar (QQp plus 259 bp (including an 86-bp intron; including 57 aa from *QQ*). Primers used were QQptarF (5'-AAAATCTGCAATTATGTAAAC-3'), QQtarRNAiR (5'-CGGGCTTCAGTTCTACAA-3'). A single construct was made for QQ RNAi: QQRNAi (a 150-bp-fragment downstream from the intron). Primers used were QQRNAiF (5'-AATTTACGTTGAAAGAAGC-3'), QQtarRNAiR (5'-CGGGCTTCAGTTCTACAA-3').

Fragments including the promoter region or promoter plus target region or *QQ* RNAi were amplified by PCR from total DNA using the attB-adapted primers (5'-GGGGACAAGTTTGTACAAAAAAGCAGGCTTC-forward primer, 5'-GGGGACCACTTTGTACAAGAAAGCTGGGTC-reverse primer). After PCR-product purification, the amplified promoter or promoter plus target region was cloned first into an entry vector pDONR221, then into a binary vector pBGWFS7 (Karimi et al., 2002; Flanders Interuniversity Institute for Biotechnology, 2003) upstream of promoterless GFP/GUS gene (Gateway™ Cloning Technology, Invitrogen, CA, 2003) or pB7GWIWG2(II) which is for hairpin RNA expression (Flanders Interuniversity Institute for Biotechnology, 2003; Karimi et al., 2002). In plants, double-stranded RNA (hairpin RNA) will be produced from the inserted sequence of interest by the RNAi construct, and post-translational gene silencing

will take place. Gateway vectors were transformed into *Agrobacterium tumefaciens* strain GV3101.

Arabidopsis Transformation and Selection

Arabidopsis plants were transformed using the floral dip method (Clough and Bent, 1998). Transformed Arabidopsis plants were selected based on Bar resistance conferred by the T-DNA. T2 seeds from at least 15 independently-transformed lines for each construct were harvested for GUS and GFP screening.

Histochemistry and Microscopy

The transgenic T2 seedlings were germinated in soil in pots. For transgenic lines containing promoter::GFP/GUS constructs, plants were harvested at various stages of development. A minimum of 3 plants were harvested for each independently-transformed line; plants or organs from the same line were stained in the same tube. Organs were stained in GUS-staining solution (Triton/Ethanol stock (Triton X-100: Ethanol: Water, 1:4:5): 0.5M PH7.0 KPO4 Buffer: 10mg/ml X-gluc in dimethyl sulfoxide: 0.1M PH7.0 Potassium ferricyanide: 0.1M PH7.0 Potassium ferrocyanide, 5:470:25:2:2) at room temperature overnight or for three h (Jefferson, 1987**)** and destained in 70% ETOH. Staining patterns were observed and documented using an Olympus stereomicroscope in the Bessey Microscopy Facility (Ames, IA).

GFP Fluorescence Microscopy

For promoter-target construct, young leaves (leaves three to six) from 15 DAI (days after planting) plants were placed in water on slides, covered with cover slips. Slides were observed and documented under a Leica TCS NT laser scanning microscope system in the Confocal Microscopy Facility (Ames, IA). Under confocal microscope, GFP signal shows green color and the auto fluorescence shows red color. The Argon/Krypton laser and 488 nm/568 nm FITC/TRITC wavelengths were used (Omnichrome, Chino, CA). We found that yellowish or damaged or dying part of the leaf would show green color under confocal

microscope. Only the healthy, green leaves were selected to detect real GFP signal. Microscope digital images were processed in Adobe Photoshop 7.0 (Adobe, San Jose, CA).

Transcriptomics Analysis

MetaOmGraph (http://www.metnetdb.org/MetNet_MetaOmGraph.htm) was used to analyze expression patterns of DBE and starch related genes. The experimental data and metadata from 70 experiments comprising 956 Affymetrix ATH1 microarray slides were obtained from two online microarray depositories: NASCArrays (http://affymetrix.arabidopsis.info/narrays/experimentbrowse.pl, Craigon et al., 2004) and PLEXdb (http://www.plexdb.org/, Shen et al., 2005). The data were normalized to the same range by scale normalization. The replicability of experiments was qualitatively assessed on scatter plots and chips with poor biological replicates were discarded. The biological replicates were averaged to yield 424 samples. This normalized data is available online (http://www.metnetdb.org/MetNet_MetaOmGraph.htm).

Reverse Transcription–PCR Analysis

Total RNA was isolated from leaves using the RNeasy Plant Mini Kit (Qiagen) and treated with DNase I (Ambion, Austin, TX). The primers used in RT-PCR are the same as QQptarF, QQpR, QQRNAiF and QQtarRNAiR. Superscript One-Step RT-PCR Kit (Gibco BRL) with 0.5 μg of total RNA per reaction was used. The amplified DNA fragments were resolved on 1.5% agarose gels with ethidium bromide.

Qualitative Comparison of Leaf Starch Content

Whole plants of QQ RNAi mutant and WT were harvested at the end of light and at the end of dark under 16 h light/8 h dark conditions. Plants were boiled in 50 mL 80% (v/v) ethanol and stained in fresh I_2/KI solution (10 g KI, 1 g I_2 in 1 L water) for 5 min, rinsed in water for 1 to 2 h. Pictures were taken immediately.

ACKNOWLEDGMENTS

We thank Xiaoli Zhang and Dr. Christophe Colleoni for providing the mutants for the microarray experiment. We would like to thank Xiaoli Zhang, Dr. Carol Foster, Dr. Jennifer Walker-Daniels, and Dr. Christophe Colleoni for helping planting and harvesting the plants for the microarray experiments.

LITERATURE CITED

Ball S, Van de Wal M, Visser R (1998) Progress in understanding the biosynthesis of amylose. Trends Plant Sci **3:** 462–467

Clough SJ, Bent AF (1998) Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. Plant J **16:** 735–743

Craigon DJ, James N, Okyere J, Higgins J, Jotham J, May S (2004) NASCArrays: a repository for microarray data generated by NASC's transcriptomics service. Nucleic Acids Res **32:** D575–D577

Crevillén P, Ballicora MA, Mérida Á, Preiss J, Romero JM (2003) The different large subunit isoforms of Arabidopsis thaliana ADP-glucose pyrophosphorylase confer distinct kinetic and regulatory properties to the heterotetrameric enzyme. J Biol Chem **278:** 28508–28515

Crevillén P, Ventriglia T, Pinto F, Orea A, Mérida Á, Romero JM (2005) Differential pattern of expression and sugar regulation of Arabidopsis thaliana ADP-glucose pyrophosphorylase-encoding genes. J Biol Chem **280:** 8143–8149

Delatte T, Trevisan M, Parker ML, Zeeman SC (2005) Arabidopsis mutants Atisa1 and Atisa2 have identical phenotypes and lack the same multimeric isoamylase, which influences the branch point distribution of amylopectin during starch synthesis. Plant J **41:** 815–830

Delatte T, Umhang M, Trevisan M, Eicke S, Thorneycroft D, Smith SM, Zeeman SC (2006) Evidence for distinct mechanisms of starch granule breakdown in plants. J Biol Chem **281:** 12050–12059

Delvallé D, Dumez S, Wattebled F, Roldan I, Planchot V, Berbezy P, Colonna P, Vyas D, Chatterjee M, Ball S, Mérida Á, D'Hulst C (2005) Soluble starch synthase I: a major

determinant for the synthesis of amylopectin in *Arabidopsis thaliana* leaves. Plant J **43:** 398–412

Dumez S, Wattebled F, Dauvillee D, Delvalle D, Planchot V, Ball SG, D'Hulst C (2006) Mutants of Arabidopsis lacking starch branching enzyme II substitute plastidial starch synthesis by cytoplasmic maltose accumulation. Plant Cell **18:** 2694–2709

Hussain H, Mant A, Seale R, Zeeman S, Hinchliffe E, Edwards A, Hylton C, Bornemann S, Smith AM, Martin C, Bustos R (2003) Three isoforms of isoamylase contribute different catalytic properties for the debranching of potato glucans. Plant Cell **15:** 133–149

Invitrogen (2003) Gateway™ Cloning Technology

Jefferson RA (1987) Assaying chimeric genes in plants: The GUS gene fusion system. Plant Mol Biol Rep **5:** 387–405

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

Kleczkowski LA (1996) Back to the drawing board: redefining starch synthesis in cereals. Trends Plant Sci **1:** 363–364

Kolbe A, Tiessen A, Schluepmann H, Paul M, Ulrich S, Geigenberger P (2005) Trehalose 6-phosphate regulates starch synthesis via posttranslational redox activation of ADP-glucose pyrophosphorylase. Proc Natl Acad Sci **102:** 11118–11123

Lunn JE, ap Rees T (1990) Apparent equilibrium constant and mass-action ratio for sucrose-phosphate synthase in seeds of Pisum sativum. Biochem J **267:** 739–743

Mentzen IW (2006) From pathway to regulon in Arabidopsis. Ph.D. dissertation. Iowa State University, Ames, IA

Moore B, Zhou L, Rolland F, Hall Q, Cheng W-H, Liu Y-X, Hwang I, Jones T, Sheen J (2003) Role of the Arabidopsis glucose sensor HXK1 in nutrient, light, and hormonal signaling. Science **300:** 332–336

Myers AM, Morell MK, James MG, Ball SG (2000) Recent progress toward understanding biosynthesis of the amylopectin crystal. Plant Physiol **122:** 989–997

Niewiadomski P, Knappe S, Geimer S, Fischer K, Schulz B, Unte US, Rosso MG, Ache P, Flugge UI, Schneider A (2005) The Arabidopsis plastidic glucose 6-

phosphate/phosphate translocator GPT1 is essential for pollen maturation and embryo sac development. Plant Cell **17:** 760–775

Niittylä T, Comparot-Moss S, Lue W-L, Messerli G, Trevisan M, Seymou MDJ, Gatehouse JA, Villadsen D, Smith SM, Chen J, et al (2006) Similar protein phosphatases control starch metabolism in plants and glycogen metabolism in mammals. J Biol Chem **281:** 11815–11818

Rook F, Corke F, Card R, Munz G, Smith C, Bevan MW (2001) Impaired sucrose-induction mutants reveal the modulation of sugar-induced starch biosynthetic gene expression by abscisic acid signalling. Plant J **26:** 421–433

Schneider A, Häusler RE, Kolukisaoglu U, Kunze R, van der Graaff E, Schwacke R, Catoni E, Desimone M, Flügge UI (2002) An *Arabidopsis thaliana* knock-out mutant of the chloroplast triose phosphate/phosphate translocator is severely compromised only when starch synthesis, but not starch mobilisation is abolished. Plant J **32:** 685–699

Sehnke PC, Chung HJ, Wu K, Ferl RJ (2001) Regulation of starch accumulation by granule-associated plant 14-3-3 proteins. Proc Natl Acad Sci USA **98:** 765–770

Sehnke PC, Henry R, Cline K, Ferl RJ (2000) Interaction of a plant 14-3-3 protein with the signal peptide of a thylakoid-targeted chloroplast precursor protein and the presence of 14-3-3 isoforms in the chloroplast stroma. Plant Physiol **122:** 235–242

Shen L, Gong J, Caldo RA, Nettleton D, Cook D, Wise RP, Dickerson JA (2005) BarleyBase—an expression profiling database for plant genomics. Nucleic Acids Res **33:** D614–D618

Smith AM, Zeeman SC, Smith SM (2005) Starch degradation**.** Annu Rev Plant Biol **56:**73–98

Stein M, Dittgen J, Sánchez-Rodriguez C, Hou B-H, Molina A, Schulze-Lefert P, Lipka V, Somerville S (2006) Arabidopsis PEN3/PDR8, an ATP binding cassette transporter, contributes to nonhost resistance to inappropriate pathogens that enter plants by direct penetration. Plant Cell **18:** 731–746

Storey JD, Tibshirani R (2003) Statistical significance for genome-wide studies. Pro Natl Acad Sci USA **100:** 9440–9445

Tang X, Ruffner HP, Scholes JD, Rolfe SA (1996) Purification and characterisation of soluble invertases from leaves of Arabidopsis thaliana. Planta **198:** 17–23

Thimm O, Blaesing O, Gibon Y, Nagel A, Meyer S, Krüger P, Selbig J, Müller LA, Rhee SY, Stitt M (2004) MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. Plant J **37:** 914–939

Tibshirani R, Walther G, Hastie T (2001) Estimating the number of clusters in a data set via the gap statistic. J R Statist Soc B **63:** 411–423

Trethewey RN, ap Rees T (1994) A mutant of Arabidopsis thaliana lacking the ability to transport glucose across the chloroplast envelope. Biochem J **301:** 449-454

Walters RG, Ibrahim DG, Horton P, Kruger NJ (2004) A Mutant of Arabidopsis Lacking the Triose-Phosphate/Phosphate Translocator Reveals Metabolic Regulation of Starch Breakdown in the Light. Plant Physiol **135:** 891–906

Wingler A, Fritzius T, Wiemken A, Boller T, Aeschbacher RA (2000) Trehalose induces the ADP-glucose pyrophosphorylase gene, *ApL3*, and starch synthesis in *Arabidopsis*. Plant Physiol **124:** 105–114

Zhang X, Myers AM, James MG (2005) Mutations affecting starch synthase III in *Arabidopsis* alter leaf starch structure and increase the rate of starch synthesis. Plant Physiol **138:** 663–674

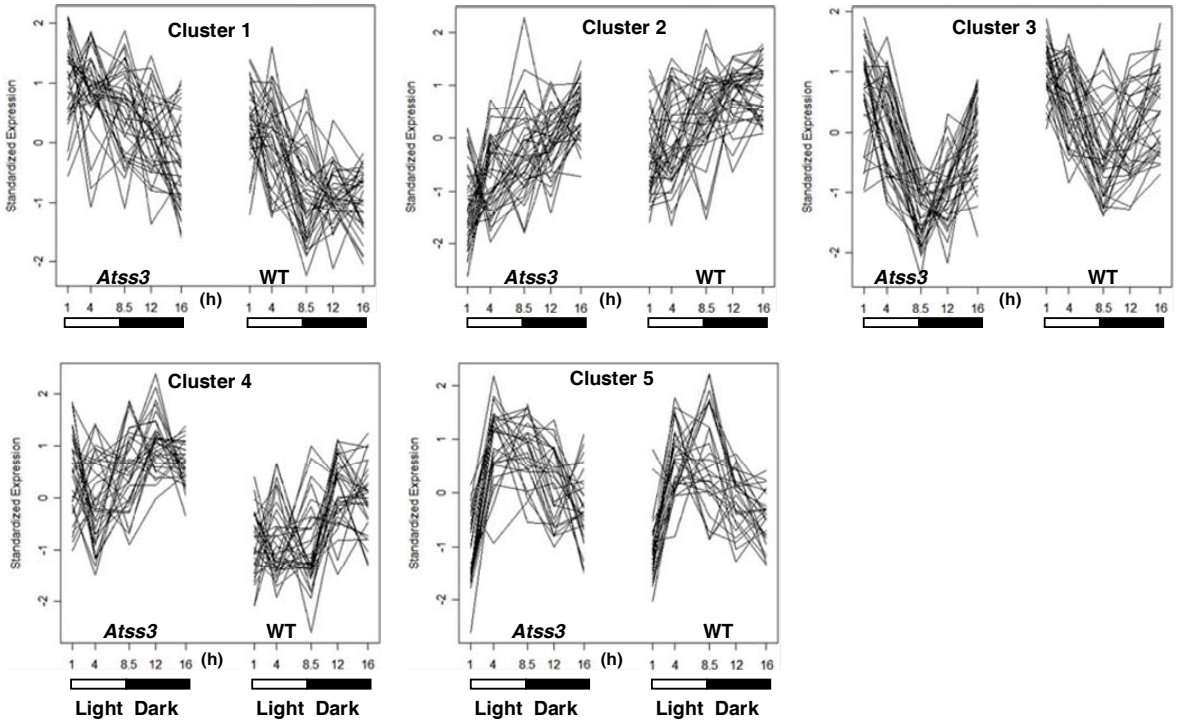


Figure 1. Clustering analysis of expression profiles of the genes that have the most significant transcript changes when FDR is controlled at the level of 0.162 in the "Genotype" comparison between *Atss3* and WT across the SD diurnal cycle.

One hundred and seventy five genes are significant when the FDR is controlled at the 0.162 level. The 175 genes are grouped into five clusters based on their expression patterns. Gene information is in Table S1 in Appendix F.

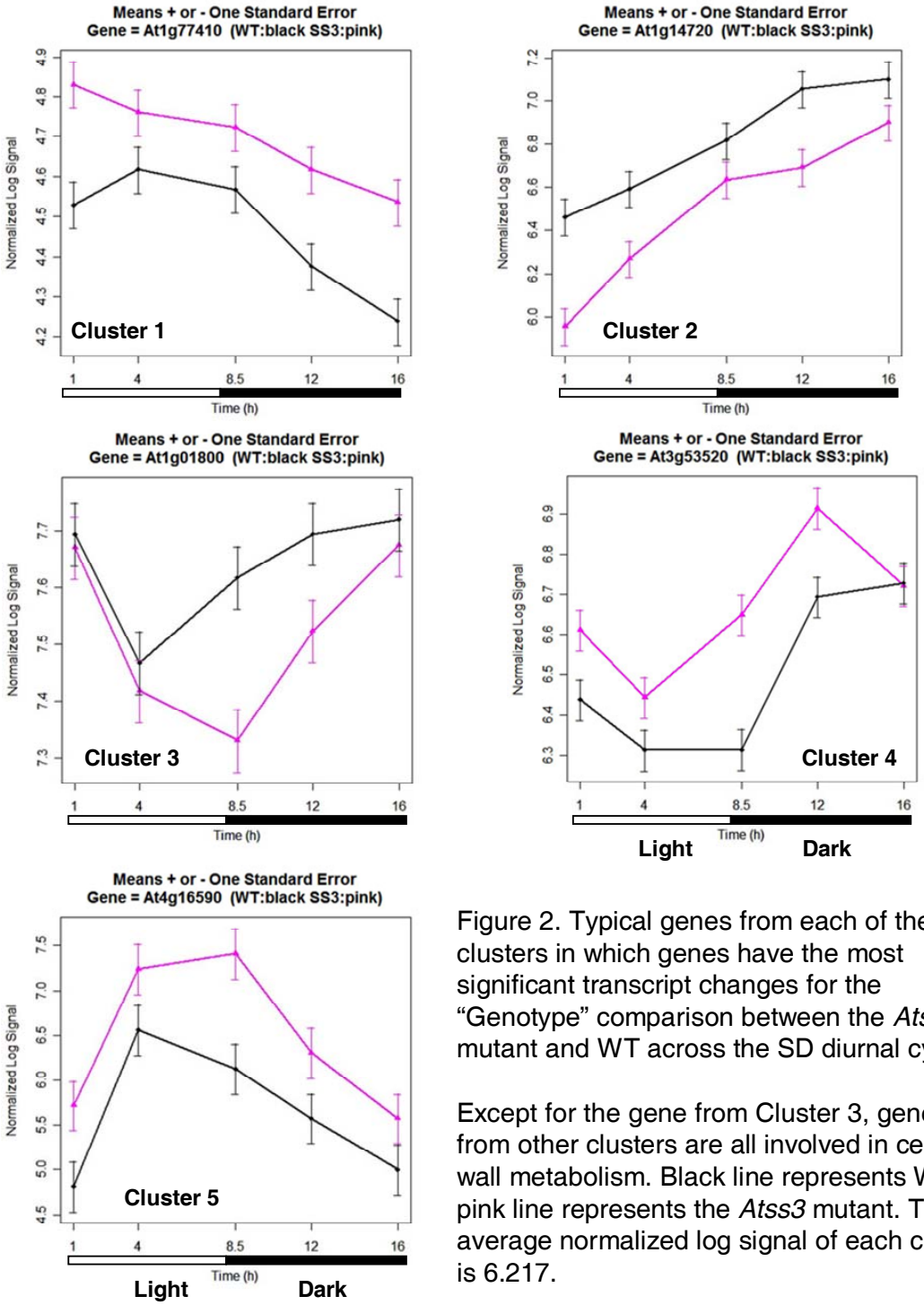


Figure 2. Typical genes from each of the five clusters in which genes have the most significant transcript changes for the "Genotype" comparison between the *Atss3* mutant and WT across the SD diurnal cycle.

Except for the gene from Cluster 3, genes from other clusters are all involved in cell wall metabolism. Black line represents WT, pink line represents the *Atss3* mutant. The average normalized log signal of each chip is 6.217.

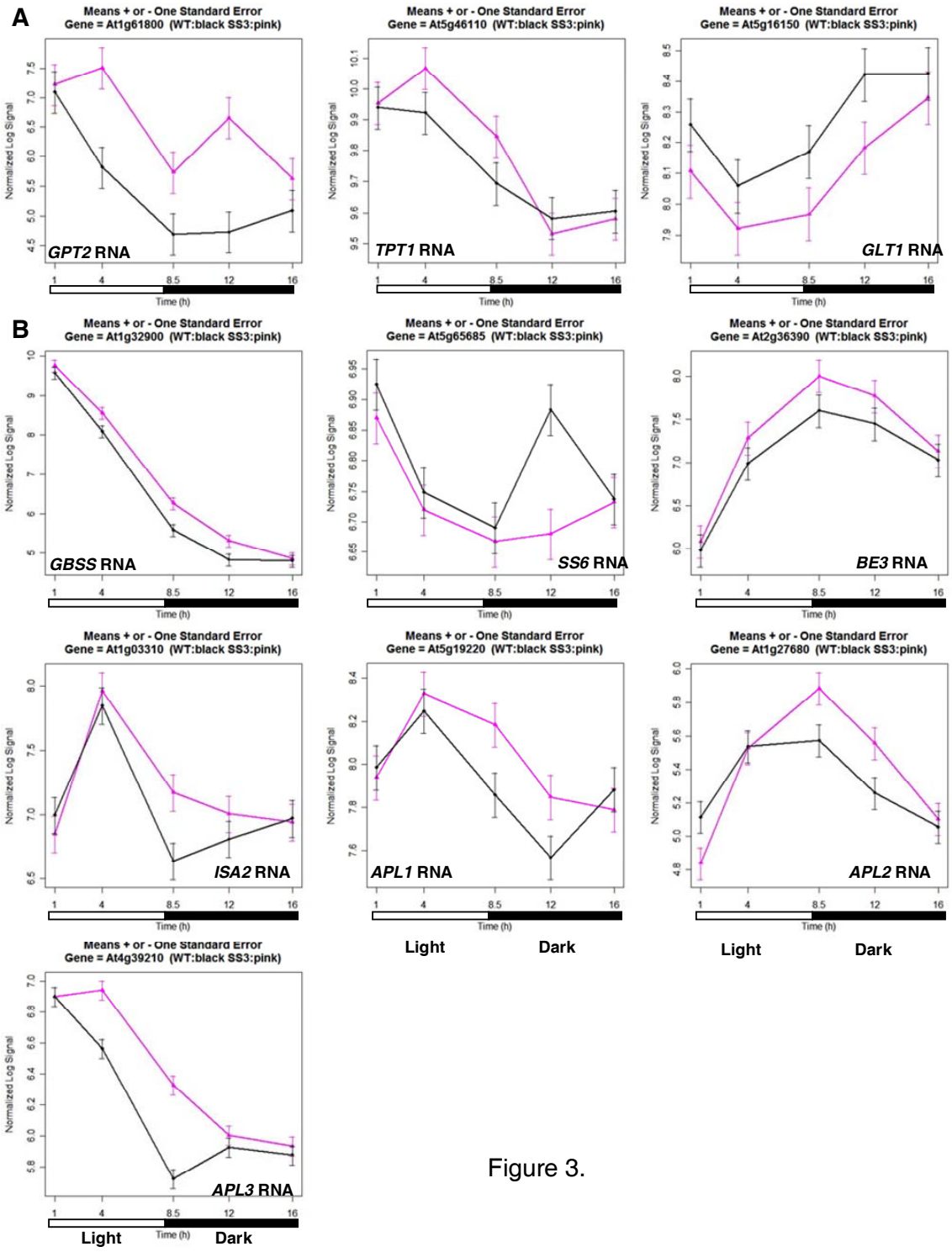

A
Means + or - One Standard Error
Gene = At1g61800 (WT:black SS3:pink)
GPT2 RNA
Means + or - One Standard Error
Gene = At5g46110 (WT:black SS3:pink)
TPT1 RNA
Means + or - One Standard Error
Gene = At5g16150 (WT:black SS3:pink)
GLT1 RNA
B
Means + or - One Standard Error
Gene = At1g32900 (WT:black SS3:pink)
GBSS RNA
Means + or - One Standard Error
Gene = At5g65685 (WT:black SS3:pink)
SS6 RNA
Means + or - One Standard Error
Gene = At2g36390 (WT:black SS3:pink)
BE3 RNA
Means + or - One Standard Error
Gene = At1g03310 (WT:black SS3:pink)
ISA2 RNA
Means + or - One Standard Error
Gene = At5g19220 (WT:black SS3:pink)
APL1 RNA
Means + or - One Standard Error
Gene = At1g27680 (WT:black SS3:pink)
APL2 RNA
Means + or - One Standard Error
Gene = At4g39210 (WT:black SS3:pink)
APL3 RNA
Normalized Log Signal
Time (h)
Light
Dark

Figure 3.

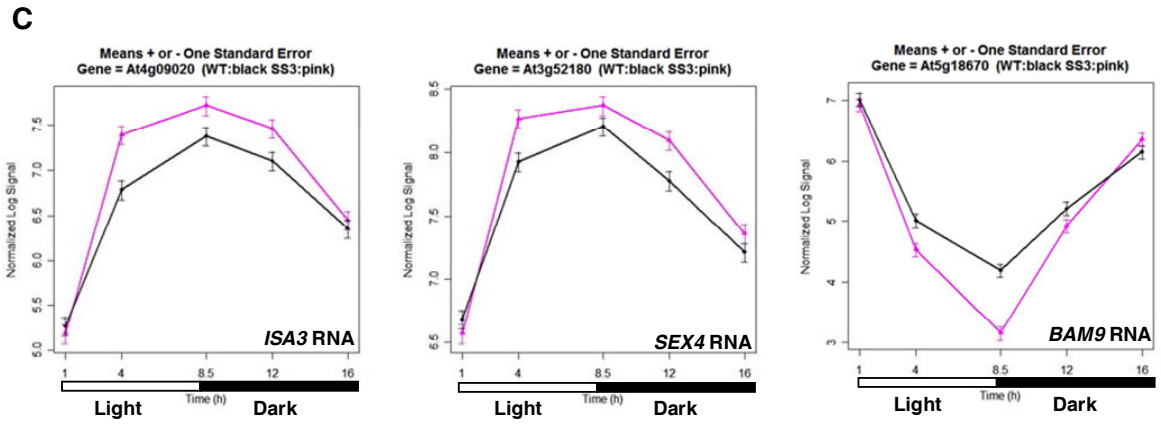


Figure 3. (continued). Genes with altered RNA accumulations related to starch metabolism in the *Atss3* mutant relative to WT.

(A) Genes involved in starch synthesis.
(B) Genes of glucose, G-6-P and triose phosphate transporter.
(C) Genes involved in starch degradation.
(D) Genes involved in sucrose metabolism.
Black line represents WT, pink line represents the *Atss3* mutant. The average normalized log signal of each chip is 6.217.

D

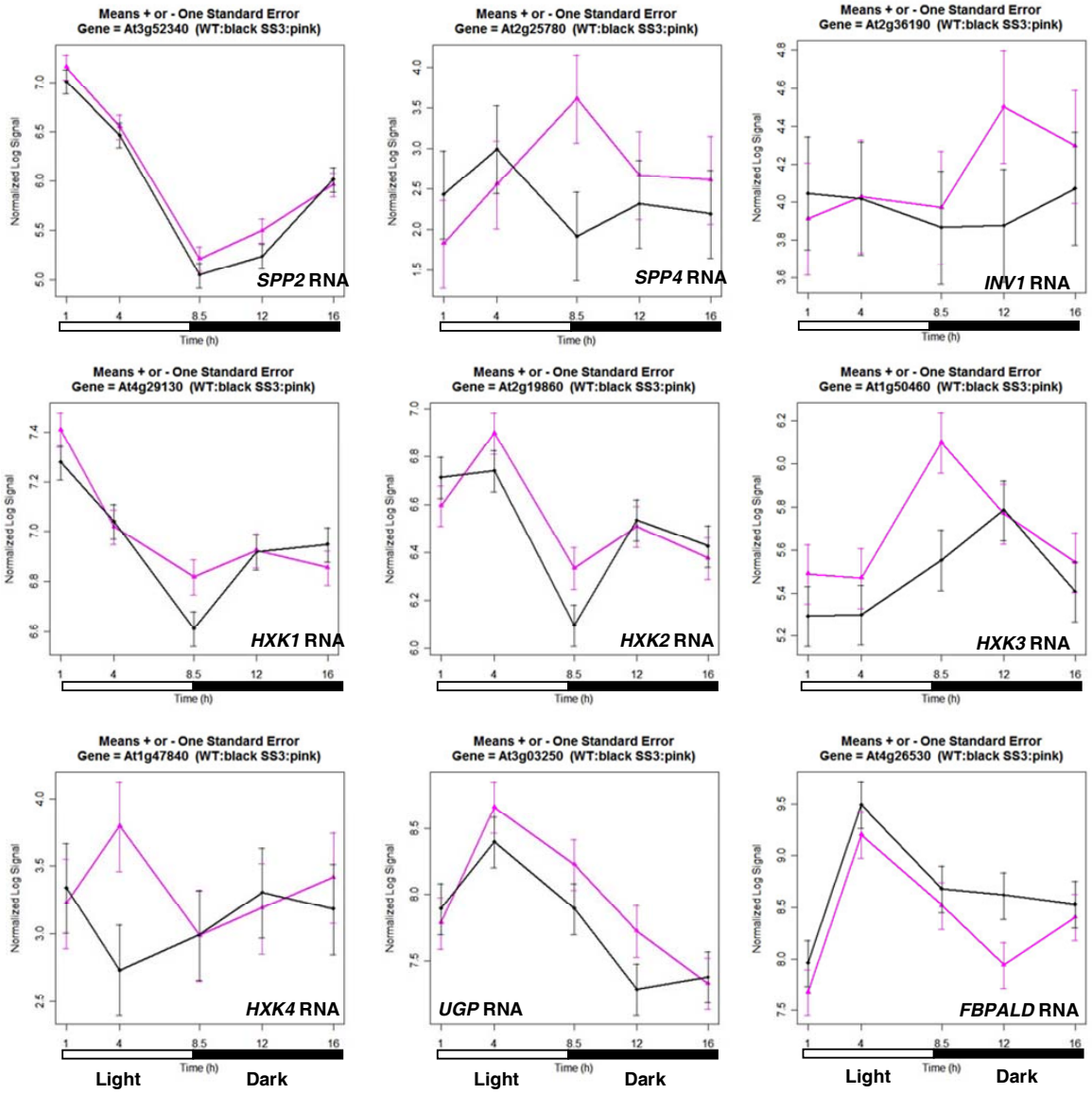


Figure 3. (continued).

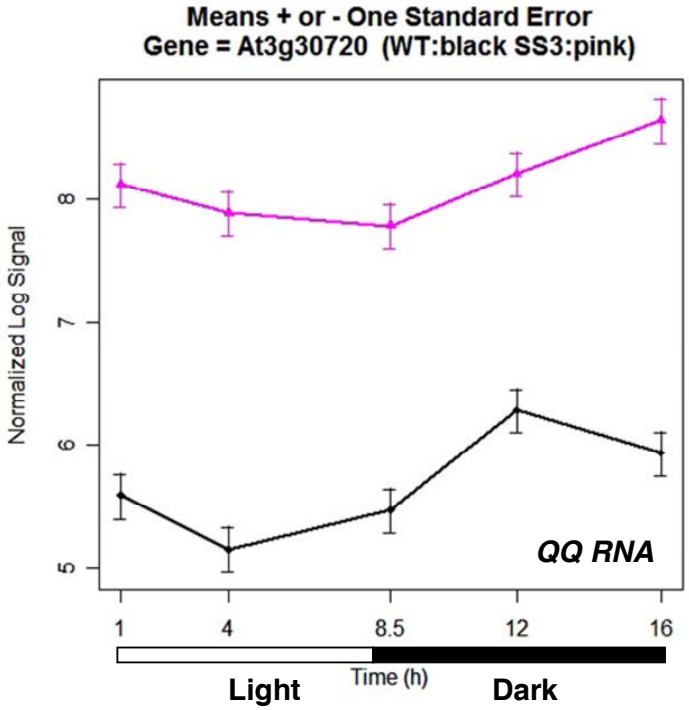


Figure 4. *QQ* is up-regulated in the *Atss3* mutant relative to WT.

Black line stands for WT, pink line stands for the *Atss3* mutant. The average normalized log signal of each chip is 6.217.

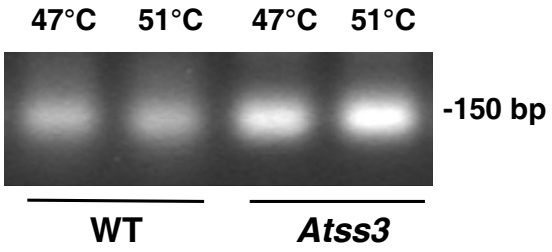


Figure 5. One-step RT-PCR confirms the *QQ* transcript is more abundant in *Atss3* mutant.

QQ RNA accumulates more in the *Atss3* mutant than in WT in samples harvested at half an hour in the dark. The reaction was repeated at two annealing temperatures: 47°C and 51°C.

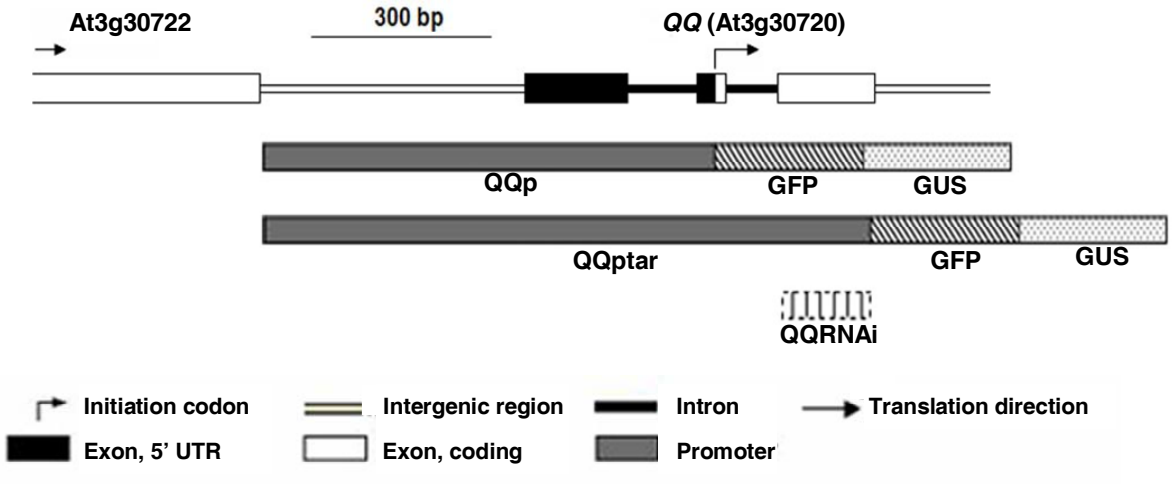


Figure 6. Structure of *QQ* gene and *QQ* promoter/promoter+target::GFP/GUS and QQ RNAi constructs.

QQp: 733 bp upstream from ATG codon (including the 114-bp intron between upstream -29 and -30 from ATG codon); QQptar: QQp plus 259 bp (including the 86-bp intron between position 18 and 19 from ATG codon, 56 aa); QQRNAi: 150 bp from position 24 to 173 from ATG codon, driven by 35S promoter.

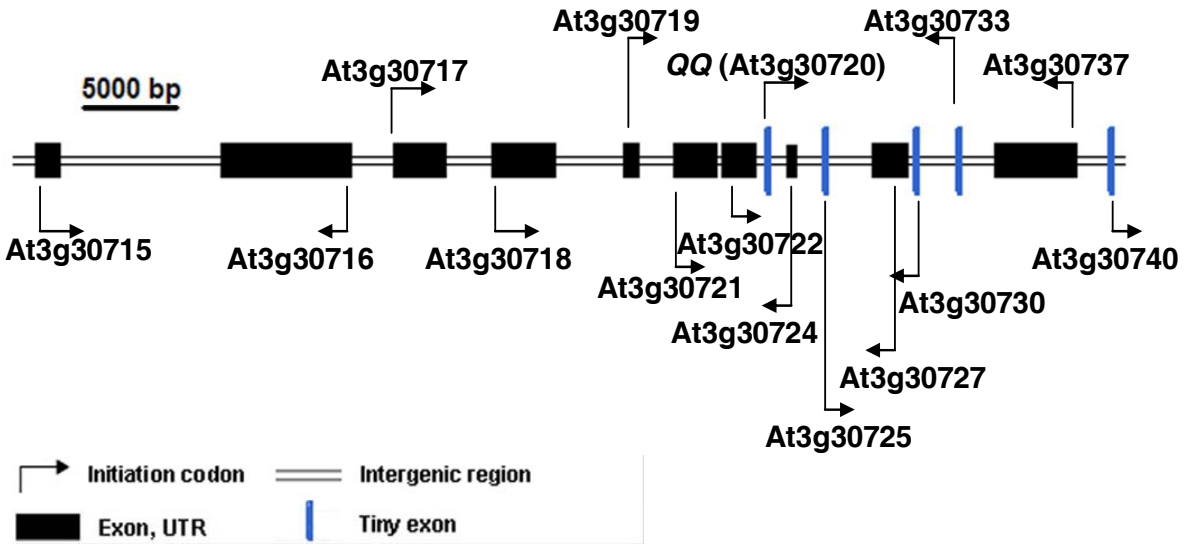


Figure 7. There are several small genes near *QQ*.

QQ and nearby genes. The tiny genes (blue) are not drawn according to their sizes.

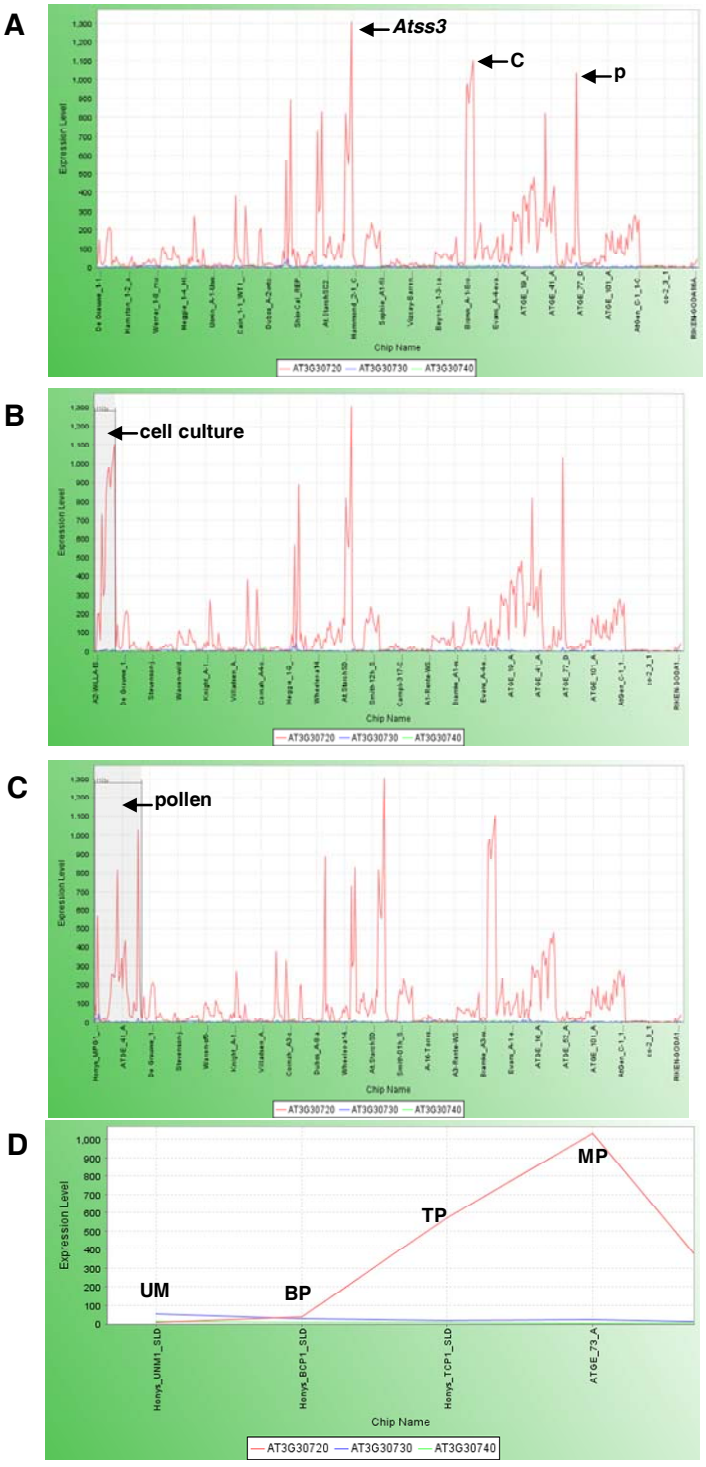

A
Atss3
C
p
Expression Level
Chip Name
AT3G30720
AT3G30730
AT3G30740
B
cell culture
C
pollen
D
MP
TP
UM
BP
Henry_UNM1_SLD
Henry_BCP1_SLD
Henry_TCP1_SLD
AT6E_73_A

Figure 8.

Figure 8. (continued). RNA accumulation of *QQ* and its nearby genes plotted in MetaOmGraph.

(A) *QQ* (At3g30720) has a complex pattern of RNA accumulation over 956 chips from public data. Two small genes (At3g30730, At3g30740) near *QQ* are not highly expressed. C=cell culture; P=pollen; *Atss3*=*Atss3* mutant.
(B) *QQ* expression in cell cultures is high.
(C) *QQ* expression in pollens is high.
(D) *QQ* RNA accumulation is increased during pollen maturation. Expression of *QQ*, At3g30730, and At3g30740 on the chips from RNA samples extracted from pollen. UM=uninucleate microspores , BP=bicellular pollen, TP=tricellular pollen, MP=mature pollen. The average expression level for each chip is 100.

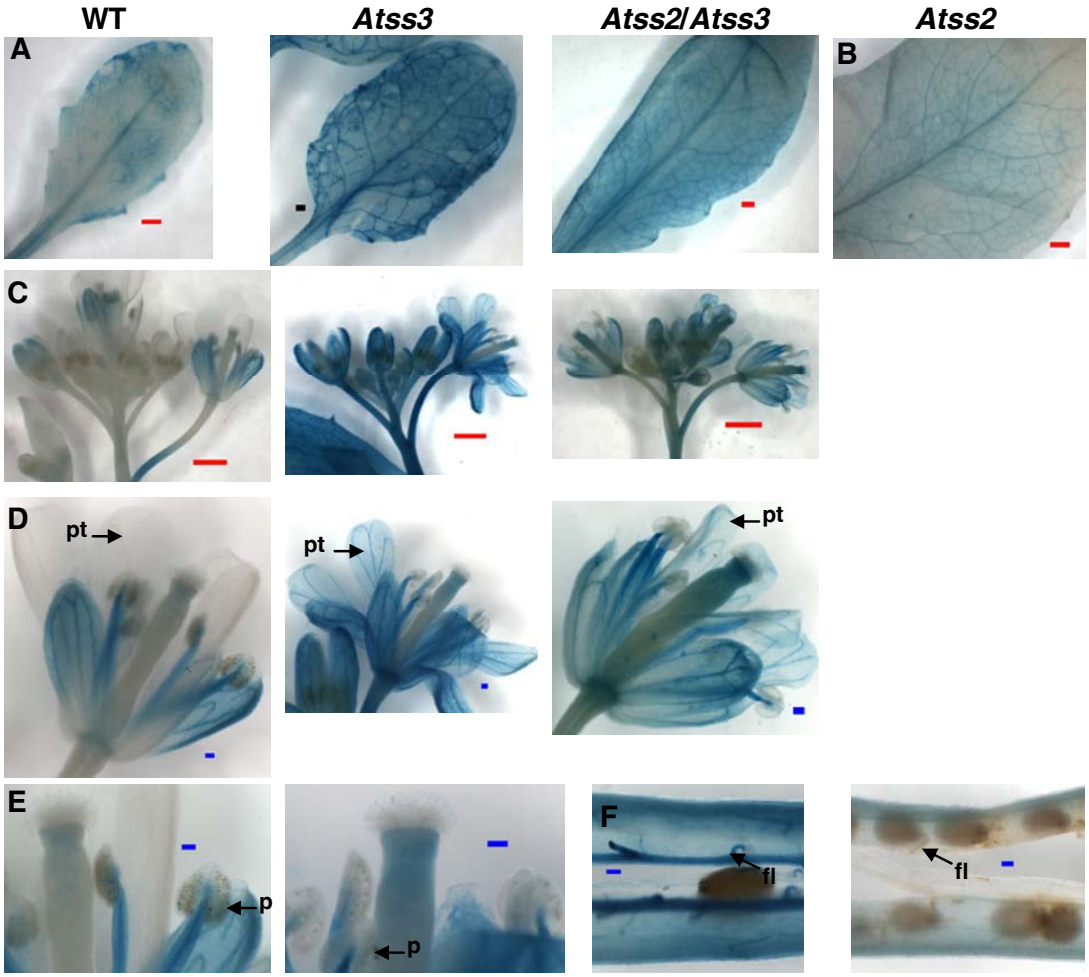


Figure 9. *QQ* is up-regulated in the *Atss3* mutant and the *Atss2/Atss3* double mutant.

The comparison of *QQ* expression in WT, *Atss3*, *Atss2/Atss3*, *Atss2* mutants: leaf (A, B), flowers and buds (C), flower (D), pollen (E), silique (F). *QQ* is lowly expressed in *Atss2* mutant in leaf (B) and silique wall (F), it is not in funicular tissue (F). *QQ* has similar expression pattern in the *Atss3* and *Atss2/Atss3* mutants. *QQ* expression in WT is similar to that in *Atss3* and *Atss2/Atss3* but at a lower level, it is in leaf (A), flowers (C, D) and pollen (E). The big difference between WT and *Atss3* and *Atss2/Atss3* mutants is that *QQ* is expressed in petal (D), and *QQ* is expressed highly in silique wall and funiculus in *Atss3* and *Atss2/Atss3* mutants (F). fl= funiculus, p=pollen, pt=petal. Red bar=1mm, blue bar=0.1mm, black bar=0.2mm.

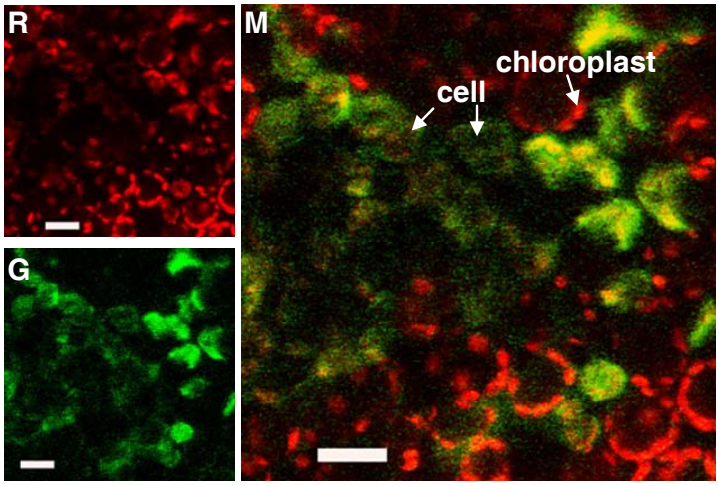


Figure 10. QQ protein appears not to be in plastids, nor nucleus.

QQ promoter-transcribed sequence-GFP fusion in leaf. The signal is patched in leaf. It is in trichome, obvious in trichome base. White bar=20μm. R represents red signal from the autofluorescence; G represents green signal from the GFP; M represents merged signal show signals from both autofluorescence and GFP.

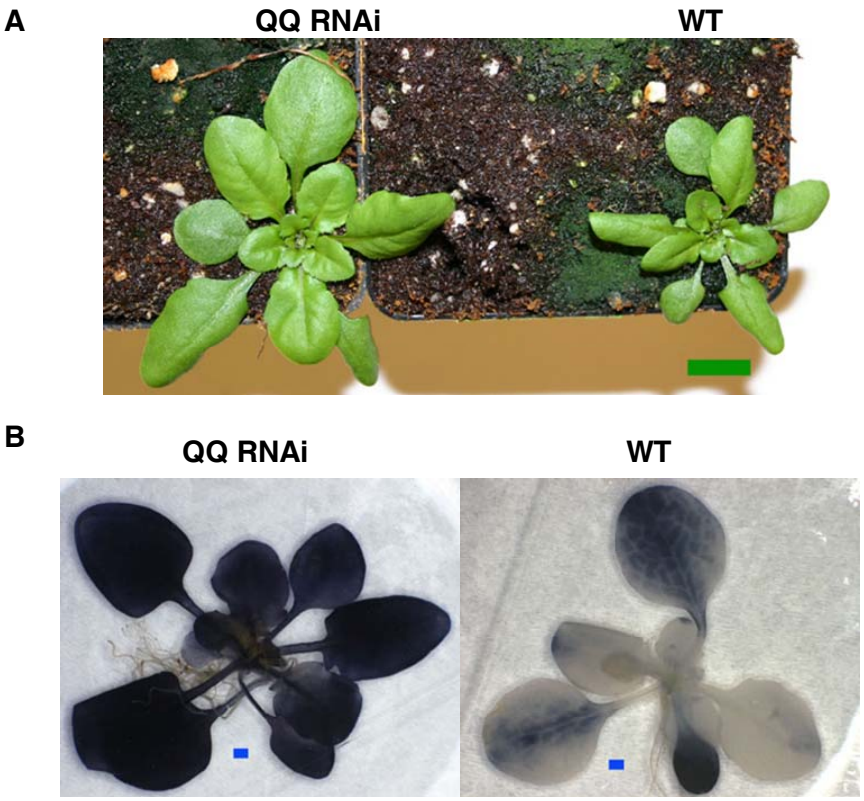


Figure 11. QQ RNAi plants grow faster than WT under LD and leaf starch of QQ RNAi seedlings indicates a starch excess at the end of light under LD conditions (16 h in light).

(A) QQ RNAi and WT plants under LD (16 h light/8 h dark).
(B) Iodine staining of the leaf starch.
Green bar=1cm, blue bar=0.1mm.

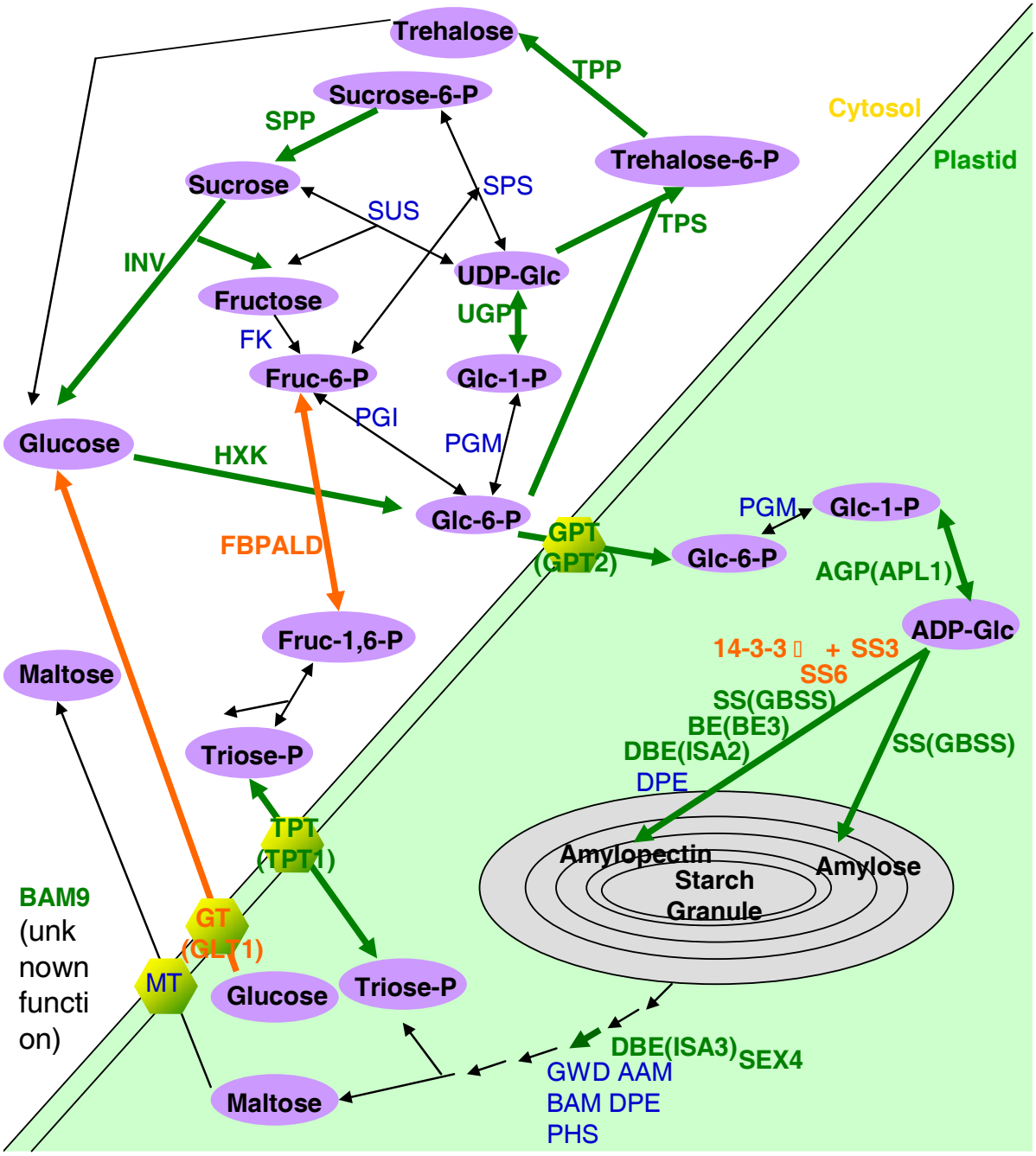


Figure 12. Genes with altered transcripts in the *Atss3* mutant in starch metabolic network.

Genes that are up-regulated in the mutant are labeled green; genes that are down-regulated in the mutant are labeled orange.

Table 1. Promoter analysis for *QQ*.

Motif	Function	Sequence	Position
CCA1 binding site motif	Light	AAAAATCT	-734
TGA1 binding site motif	bZip transcription factor	TGACGTGG	-504
UPRMOTIFIAT	Unfolded protein response	TGACGTGG	-504
ABRE-like binding site motif	ABA	GACGTGGC	-503
ACGTABREMOTIFA2OSEM	ABA	ACGTGGC	-502

Table 2. Little is known about *QQ* and its nearby genes. Some of them (including *QQ*) are quite small genes with unknown function.

Locus_ID	cDNA (bp)	Predicted protein (aa)	Predicted direction of transcription relative to QQ	Tair_annotation	Cellular component	Represented in ATH1 chip
AT3G30715	1279	-	Forward	expressed protein		No
AT3G30716	6998	-	Reverse	gypsy-like retrotransposon family		No
AT3G30717	2799	-	Forward	gypsy-like retrotransposon family		No
AT3G30718	3401	-	Forward	gypsy-like retrotransposon family		No
AT3G30719	855	-	Forward	CACTA-like transposase family		No
AT3G30721	2337	-	Forward	CACTA-like transposase family		No
AT3G30722	1785	-	Forward	CACTA-like transposase family		No
AT3G30720	380	60	Forward	expressed protein	unknown	Yes
AT3G30724	509	-	Reverse	hypothetical protein, non-LTR retroelement reverse transcriptase		No
AT3G30725	336	112	Forward	expressed protein	endomembrane	No
AT3G30727	1841	-	Reverse	non-LTR retrotransposon family		No
AT3G30730	204	68	Reverse	hypothetical protein	unknown	Yes
AT3G30733	296	-	Reverse	Hypothetical protein		No
AT3G30737	4407	-	Reverse	putative helicase		No
AT3G30740	217	-	Forward	40S ribosomal protein S25		Yes

QQ is marked with yellow. *QQ* has no nucleotide or protein sequence similarity to any gene of Arabidopsis or other organisms with sequenced genomes.

Table 3. Genes with RNA accumulation profiles highly correlated with that of *QQ*.

Corr	Locus ID	Tair annotation
1.00	AT3G30720	expressed protein
0.60	AT3G47120	RNA recognition motif (RRM)-containing protein
0.56	AT4G37295	expressed protein
0.54	AT2G03830	expressed protein
0.54	AT2G23060	GCN5-related N-acetyltransferase (GNAT) family protein
0.53	AT1G33680	KH domain-containing protein
0.53	AT5G59060	expressed protein
0.52	AT2G47520	a member of the ERF (ethylene response factor) subfamily B-2 of ERF/AP2 transcription factor family
0.52	AT5G18190	protein kinase family protein
0.52	AT5G39820	no apical meristem (NAM) family protein
0.52	AT1G51190	AINTEGUMENTA-like (AIL) subclass of the AP2/EREBP family of transcription factors
0.52	AT3G45770	oxidoreductase, zinc-binding dehydrogenase family protein
0.51	AT3G19200	expressed protein
0.51	AT2G14960	IAA-amido synthases, insertion in this gene are hypersensitive to auxin.
0.51	AT3G28970	expressed protein
0.51	AT4G32890	GATA transcription factor 3
0.51	AT3G60650	expressed protein
0.50	AT1G05270	TraB family protein
0.50	AT3G10780	emp24/gp25L/p24 family protein
0.50	AT5G26870	MADS-box family protein
0.50	AT5G10040	expressed protein
0.50	AT5G50550 AT5G50650	WD-40 repeat family protein / St12p protein

Genes that are 0.5 and above correlated with *QQ* (labeled yellow) in MetaOmGraph: about 1/3 of them are unknown “expressed protein”; nearly 1/3 of them are involved in RNA recognition or are transcription factors (labeled grey). “Corr” stands for correlation coefficient.

CHAPTER 5. GENERAL CONCLUSIONS

The data in the body of this dissertation provide a greater understanding of starch metabolism and its regulation. Using mutants altered in starch accumulation, a combination of approaches such as global gene profiling, and promoter::GUS/GFP analyses and genetics, I have expanded the understanding of starch metabolism.

Starch DBEs hydrolyze α-(1→6) linkages and are critical for accumulation of normal levels, morphology, and composition of starch granules. The functions and the relations among the members of the DBE gene family are not yet understood. In Chapter 2, I explored the relationship among the DBE family members using a combination of promoter::GUS/GFP and bioinformatics analyses. These studies based on co-expression of *ISA1* and *ISA2*, and promoter motifs, are consistent with previous evidence that ISA1 and ISA2 functioning together as a protein complex (ISA1/ISA2) for starch synthesis. They also show that *ISA2* and *ISA3* are highly co-expressed, both spatially and temporally, and the promoters of these have several putative regulatory motifs in common. This suggests these polypeptides might also form a protein complex, taking part in starch metabolism. My research also gives experimental evidence that ISA3 protein is localized in the leaf chloroplast, consistent with the possibility of an ISA2/ISA3 complex. These studies lead to a testable hypothesis about the co-function of ISA2 and ISA3. Based on patterns of expression of the DBE genes in response to stress and analysis of the promoters, I hypothesize that ISA2/ISA3 may play a role in starch metabolism in response to moisture and cold acclimation.

Starch biosynthesis and degradation are diurnally-regulated processes with many enzymes involved. In Chapter 3, I describe seven clusters of co-expressed genes in WT plants grown under a SD diurnal cycle. Different expression patterns might represent the typical processes over the diurnal cycle. In Chapter 3 and 4, I compare the changes in gene expression associated with mutants that accumulate increased levels of starch.

Two very different Arabidopsis mutants, the antisense-*ACLA* mutant (which has a reduced level of ATP citrate lyase (ACL), the enzyme catalyzes the synthesis of cytosolic acetyl-CoA), and the *Atss3* T-DNA knockout mutant (loss-of-function of the SS3 starch

synthase), both have a starch-excess phenotype. Changes in gene transcript accumulation identified by mRNA profiling comparison indicate a set of starch metabolic genes that may contribute to this phenotype. Specifically, both mutants have in common a subset of up-regulated genes of starch synthesis: ADP-glucose pyrophosphorylase (AGPase) regulatory subunits (*APL3* and *4* in the antisense-*ACLA* mutant and *APL1-3* in the *Atss3* mutant); glucose-6-phosphate/phosphate translocator (*GPT1* and *2* in the antisense-*ACLA* mutant and *GPT2* in the *Atss3* mutant, in transporting starch synthesis precursor into the plastid); and trehalose synthesis (*TPP* in the antisense-*ACLA* mutant, *TPS* and *TPP* in the *Atss3* mutant, in trehalose biosynthesis). Trehalose is thought to stimulate starch synthesis via activation of AGPase. Both mutants have down-regulated beta-amylase-like genes (*BAM1* in antisense-*ACLA* mutant and *BAM9* in the *Atss3* mutant).

In addition, both mutants have changes in levels of transcripts of genes involved in cell wall metabolism. This might reflect the carbon partitioning into starch by modification of carbon into cell wall.

Acetyl-CoA plays a critical role for many biochemical pathways essential to plant growth and development. A decreased cytosolic ACL might result in insufficient cytosolic acetyl-CoA substrate for generation of the isoprenoids and polyketides derived from this molecule. This alteration could affect the plant signaling. Specifically, in these antisense-*ACLA* mutants, the APX1 pathway genes are down-regulated. The down-regulated APX1 pathway is thought to induce oxidative stress. It is possible that redox-regulation of starch metabolism results in a carbon partitioning into starch. These processes could affect and regulate each other. They represent testable hypotheses that could elucidate metabolic interconnections between starch accumulation and acetyl-CoA metabolism, two disparate metabolic processes.

The loss-of-function of a SS3 (starch synthase) was previously shown to cause a starch-excess phenotype despite having a very normal opposing plants. My data indicate that many genes involved in starch synthesis are up-regulated in this mutant. Increased *GPT2*, *TPT1*, *ISA2*, *GBSS*, *BE3* and *APL1-3* transcripts might lead to the increased starch synthesis. The increased transcripts of *GBSS*, *ISA2* and *ISA3* might explain the alteration in the amylopectin structure. Also, the up-regulated genes in sucrose metabolism (*SPP2*/*4*, *INV1*,

HXK1-4 and *UGP*) in the mutant may indicate a direction of change toward an increase of G-6-P, which could be transported to chloroplast as a precursor for starch biosynthesis. Metabolic flux studies in this mutant might give more information about the starch metabolic network.

My hypotheses of the possible changes in processes of the starch metabolic network are based on the microarray data at the transcription level. What happens in the plant regarding to the post-transcriptional, translational and post-translational level is still unknown. Also, microarray data do not provide evidence for caused relationship between genes. In the course of these studies, I identified several genes of unknown function that were altered in expression in the antisense-*ACLA* or *Atss3* mutants. In the *Atss3* mutant, a gene (referred to as *QQ*) is increased by more than 7-fold across all time points. I decided to focus on this gene experimentally to determine if it has a role in starch metabolism. *QQ* encodes a tiny protein of only 59 aa. It is also up-regulated in the *Atss2*/*Atss3* double mutant across a variety of developmental time points, in leaves, petals and funiculus. I generated QQ RNAi lines to test whether they had an altered starch phenotype. QQ RNAi mutant lines with reduced *QQ* transcript have a WT appearance but contain more starch than do WT. This supports the involvement of QQ in starch metabolism. Since QQ is predicted to be a so tiny protein with no known catalytic motifs, it may be more likely that QQ has a regulatory role rather than a catalytic function.

APPENDIX A. EXPRESSION OF GENES RELATED TO THE ACETYL-COA/BIOTIN NETWORK

CAC1A and *CAC1B*

ACCase catalyzes the carboxylation of acetyl-CoA to produce malonyl-CoA: HCO_3^- + ATP + acetyl-CoA → malonyl-CoA + ADP + Pi

Malonyl-CoA in plastids is the direct precursor for *de novo* fatty acid synthesis. Malonyl-CoA in the cytosol is used for fatty acid elongation and secondary metabolites synthesis such as cuticular waxes. There are two forms of ACCase in plants: heteromeric (ht)ACCase in plastid and homomeric (hm)ACCase in cytosol (Konishi and Sasaki, 1994; Alban et al., 1994). Biotin carboxyl carrier protein (BCCP), biotin carboxylase (BC) and biotin carboxyltransferase (CT) are three domains of ACCase. These three domains are located on one polypeptide for hmACCase and on different polypeptides for htACCase. The heteromeric acetyl-CoA carboxylase (htACCase) catalyzes the first reaction of *de novo* fatty acid biosynthesis in plastids. HtACCase consists of 4 subunits: nuclear encoded BCCP, BC, α-CT, and plastid-encoded β-CT. *CAC1* encodes the BCCP (Choi et al., 1995; Ke et al., 1997; Ke et al., 2000). BCCP has two isoforms: BCCP1 and BCCP2. BCCP1 is encoded by *CAC1A* (At5g16390). It is present in developing tissues such as roots, leaves, stems, flowers, siliques and seeds. BCCP1 protein is present in all developing organs. BCCP2 is encoded by *CAC1B* (At5g15530). It has 42% amino acid identity to BCCP1. BCCP2 transcript is present in flowers and siliques, and not in leaves, roots or stems (Thelen et al., 2001).

Previous studies in our lab (Qian, 2002) show that antisense-*CAC1A* transgenic plants contain reduced BCCP1; expression of other ACCase genes is not affected. The antisense-*CAC1A* plants with reduced BCCP1 show unique morphological changes: smaller plant, leaf and cell size, crinkly leaf, yellow patches or variegated yellow pattern. The severity of mutant phenotype of antisense plants is correlated with the reduction of BCCP1. In antisense-*CAC1A* plants, fatty acid content per weight of leaf or seed is not affected; but is reduced per plant. Fatty acid composition in leaf and seeds of antisense plants is similar to that of WT plants. So BCCP1 is important for fatty acid synthesis. Lack of BCCP1 will affect plant's metabolism.

Since the *CAC1A* and *CAC1B* probes used for *in situ* hybridization are 73% identitical (152/206) to each other, a northern blot was done to check whether the *CAC1A* probe and *CAC1B* probes cross-hybridize. Two RNA gels were loaded with the same amount of *CAC1A* and *CAC1B* sense RNA, and hybridized separately with *CAC1A* and *CAC1B* antisense probe. Figure 1A shows the probes do not cross-hybridize.

The antisense-*CAC1A* plants have reduced level of BCCP1 (Qian, 2002). Figure 1B shows only *CAC1A* mRNA is distributed in leaves and developing siliques of both wild type and antisense-*cac1a* mutant. In contrast, *CAC1B* mRNA accumulates in developing siliques of both wild type and antisense-*cac1a* mutant but is not detected in leaves. *CAC1A* and *CAC1B* have different expression patterns in leaves. *CAC1A* RNA is present in the antisense-*cac1a* mutant. The antisense-*cac1a* mutant does not affect the accumulation of *CAC1B* RNA in siliques. There is no difference in the accumulation of *CAC1A* mRNA between antisense-*CAC1A* plants and WT. One possibility for the lack of a *CAC1A* mRNA reduction is that in antisense-*CAC1A* plants, *CAC1A* is not affected at the transcription level, but at the translation or post-transcriptional level.

To compare the cellular distribution of *CAC1A* and *CAC1B* mRNAs within embryos and siliques, particularly at the earlier stages when embryos cannot be dissected from the siliques, relative accumulation of these mRNAs was determined by *in situ* hybridization at 5 stages starting from 1 DAF (Figure 2). In rapidly enlarging siliques (1 DAF), the sporophyte within the ovule is a zygote or several-celled embryo, and *CAC1A* mRNA is evenly distributed throughout the silique, including the silique walls and developing ovules. By 3 DAF, siliques have ceased enlargement. *CAC1A* mRNA is concentrated in the rapidly growing globular embryos, and is dramatically reduced in silique walls and ovule integuments. At 5 to 7 DAF, the time of maximal seed oil accumulation, *CAC1A* mRNA is highly concentrated in the embryos. By 12 DAF, embryos have ceased growing and are approaching desiccation; at this stage, *CAC1A* mRNA is no longer detected in the embryos or siliques.

Like *CAC1A*, *CAC1B* mRNA is distributed evenly throughout the silique tissues at 1 DAF. By 3 DAF, *CAC1B* mRNA is concentrated in the embryos and is reduced in silique walls and integuments of the ovules. In the 5 DAF and 7 DAF siliques, *CAC1B* mRNA is

concentrated in the embryos. In the 12 DAF siliques, signal is no longer detected in the embryos.

In situ hybridization showed that *CAC1A* and *CAC1B* genes were expressed with similar spatial and temporal patterns during embryo development. These observations suggest that BCCP1 is important for htACCase activity *in planta*, but BCCP2 is dispensable.

To better understand the expression and function of *CAC1A* and *CAC1B*, we investigated their cellular localization. We constructed vectors that expressed fused GFP/GUS under the control of the *CAC1A* and *CAC1B* gene promoter regions (Figure 3). The upstream intergenic region is 701 bp for *CAC1A* and 808 bp for *CAC1B*. Usually the promoter for Arabidopsis is within between 1000 bp and 1500 bp. Therefore, we made three constructs for *CAC1A*: CAC1Ap1 (1124 bp upstream of ATG codon, obtained from Dr. Nikolau lab), CAC1Ap2 (884 bp upstream of ATG codon), and CAC1Ap3 (701 bp upstream of ATG codon). We made two constructs for *CAC1B*: CAC1Bp1 (1244 bp upstream of -13 from ATG codon, obtained from Dr. Nikolau lab) and CAC1Bp2 (807 bp upstream of ATG codon). The constructs from same gene gave the similar results. The promoter of *CAC1A* is within 701 bp upstream of ATG codon and the promoter of *CAC1B* is within 807 bp upstream of ATG codon.

Analysis of promoter-driven GUS expression (Figure 4) shows *CAC1A* is expressed in cotyledons, leaves (leaf vasculature and leaf mesophyll cell) and shoot meristem from young seedling stage to post-bolting. It is expressed in flowers in sepal, petal, filament, stigma, young silique, young embryo, receptacle, and pedicle. It is expressed in stem and trichomes. It is not expressed in anther, pollen, or old silique. It is highly expressed in roots except the root tip. *CAC1B* is not detectable in seedling. In contrast, *CAC1B* is detected in flowers, sepals, filaments, stigma, and young siliques, young embryos and receptacles. It has a similar expression pattern in roots as *CAC1A* does but with reduced intensity (Table 1). Where there is *CAC1B* expression, there is much higher expression of *CAC1A*. The expression in silique and embryo is consistent with *in situ* result (Figure 2). The GUS expression pattern also supports the hypothesis that BCCP1 is important for htACCase activity *in planta*, but BCCP2 is dispensable.

CACT

At5g46800 is Carnitine/acylcarnitine translocase-like protein. CACT has been reported in plants (Lawand et al., 2002). However, the data of Fatland et al. (2005) indicate acetyl-CoA is not transported from mitochondrion to cytosol. Many people doubt the existences of CAT/CACT in the plant. To begin to understand where in the plant this gene is expressed, I did *in situ* in different stages of *Arabidopsis* siliques. The *in situ* results (Figure 5) show that *CACT* is expressed in *Arabidopsis* siliques. *CACT* mRNA is distributed all over the silique tissues at 1 DAF. It is reduced in the silique walls and mainly concentrated in the embryo at 3 DAF. From 5 DAF to 7 DAF, it is highly concentrated in the embryo, reaching peak at 5 DAF. It decreases after 7DAF and almost non-detectable at 12 DAF. This pattern of gene expression is consistent with a role in embryogenesis.

To better understand the expression and function of CACT, we investigated its cellular localization. We constructed vectors that expressed fused GFP/GUS under the control of the *CACT* promoter (Figure 6). There is a 682-bp intron in the *CACT* 5' UTR from upstream -3 of ATG codon. So we chose the 1340 bp upstream of ATG codon as the promoter region.

CACT expression pattern is similar to that of *CAC1A*. Figure 7 shows it is expressed in cotyledons, leaves (leaf vasculature and leaf mesophyll cell) and shoot meristem from young seedling to mature plant. It is in flowers in sepal, filament, stigma, young silique, young embryo, receptacle, and pedicle. It is in stem and trichomes. It is not in anther, and not in old silique. It is highly expressed in roots. How it differs from *CAC1A* is that it is expressed in pollen but not in petal. It is expressed in root tip. Table 1 includes a summary of *CACT* expression pattern.

To investigate the sub-cellular localization of CACT, we made constructs that expressed GFP/GUS fused to N-terminal of CACT, under the control of the *CACT* promoter region. The predicted cleavage site is at 22 aa. CACT is only 300 aa. We included the whole protein for the promoter plus target construct (Figure 6). Figure 8 shows clearly CACT is not localized in chloroplast. The method limits us from further identification of its sub-cellular localization.

Wax Synthase

Waxes are oxygen esters of primary fatty alcohols and fatty acids. Three key enzymatic activities are required for wax biosynthesis: fatty acyl-CoA elongase (FAE), fatty acyl-CoA reductase (FAR) and wax synthase (WS). FAE system supplies very-long-chain acyl-CoA, the precursors for wax biosynthesis (Lassner et al., 1996). FAR carries out the reduction of acyl-CoA to yield fatty alcohols (Metz et al., 2000). Wax synthase (WS) is fatty acyl-CoA: fatty alcohol acyltransferase. It catalyzes the final step of the transfer of an acyl chain from fatty acyl-CoA to a fatty alcohol in the synthesis of linear esters (waxes) (Kahn and Kolattukudy, 1973; Kolattukudy and Rogers, 1978, 1986; Wu et al., 1981). Conventional oilseed crops produce triacylglycerols (TAG) as the primary storage lipid, such as soybean and corn. However, Jojoba seed contains a wax ester as the primary storage lipid (Lardizabal et al., 2000). Up 50% of the Jojoba seed's dry weight is wax. On Arabidopsis chromosome 5, there are 8 genes with homology to *WS* (Lardizabal et al., 2000). Seven of them are next to each other. Which of these, if any, actually functions as WS, is not yet known. It is possible the WS-like genes have a rele in plant development and they might function in the formation of seed lipids.

There are 12 genes in the Arabidopsis genome encoding wax synthase-like protein family members: At5g51420, At5g55320, At5g55330, At5g55340, At5g55350, At5g55360, At5g55370, At5g55380, At1g34490, At1g34500, At1g34520, and At3g51970. Several of these are in physical proximity: seven in Chromosome 5 and 3 in Chromosome 1. We are very interested in their functions and the relationship among them. We chose two genes located in chromosome 5: At3g55380 (*WaxS*) and At5g55340 (*WSlike*), to study.

I constructed vectors that expressed fused GFP/GUS under the control of the wax synthase genes' promoters (Figure 9). The upstream intergenic region for *WaxS* is only 280 bp. I made two constructs: WaxSp1 (1551 bp upstream of ATG codon) and WaxSp2 (280 bp upstream of ATG codon). The upstream intergenic region for *WSlike* is 1330 bp. We made a single construct for *WSlike*: WSlikep (1330 bp upstream of ATG codon).

WaxSp1 and WaxSp2 expression patterns are indistinguishable from each other. The effective promoter of the expression that directs *WaxS* during developing is within 280 bp upstream of start codon.

Figure 10 shows *WaxS* has strong expression in cotyledons, and true leaves, in particularly mesophyll cells and vasculature. It is expressed in almost all vasculatures of the shoot. It is expressed in shoot base and petiole, but not in shoot meristem. *WaxS* is expressed strongly in roots, but except developing de novo formation of lateral roots, or in the root tip (except for the cells of the root cap). It is expressed in root vasculature and less in cortex.

Just before bolting, it shows strong expression in older leaves, but much less in young expanding leaves, it is not expressed in shoot bolting buds. High expression is maintained in roots, vasculatures, and root caps.

WaxS has low expression in stems of bolting plant and strong expression in older rosette leaves. The expression in cauline leaves is reduced relative to that of rosette leaves. It has strong expression in opened flower and young buds in sepals, and pedicles of opened flowers, and also in the filament except the top part, and mature pollen. It is not detectable in petal, anther, or receptacle. It is expressed at a lower level in the young and old silique wall. It is expressed in cauline leaf trichome base.

WSlike is not expressed in cotyledon. It is expressed in shoot meristem and young leaves. Just before bolting, it is expressed in shoot meristem and young leaves. It is expressed in leaf mesophyll cells, not obvious in veins. It is expressed in bolting buds. After bolting, it has very weak expression in shoot meristem and young leaves. It is concentrated at stems just under flowers, cauline leaf and stems, in meristem around the cauline leaf. It is expressed in receptacle and filament top part. It is not expressed in stigma, nor in pollen. It is expressed in cauline leaf trichome top. It is not expressed in silique. It is extraordinarily expressed in root tip, in germinating roots. It is not expressed all over the roots: it is expressed in root epidermal cells, root nodes surface. It is expressed in secondary roots (Figure 10). The summary of the expression patterns of *WaxS* and *WSlike* is in Table 1.

The relation of the expressions of *WaxS* and *WSlike* is different from the expressions of *CAC1A* and *CAC1B*. Unlike *CAC1A* expression overlaps the expression of *CAC1B*, the expression of *WaxS* and *WSlike* is complementary. Expression of *WSlike* is kind of opposite to *WaxS*. *WSlike* is usually expressed in where there is no *WaxS* expression. *WaxS* and *WSlike* have different temporal and spatial expression patterns. They might have different functions.

To investigate the sub-cellular localization of WaxS and WSlike, we made constructs that expressed GFP/GUS fused to N-terminal of WaxS/WSlike, under the control of the *WaxS*/*WSlike* promoter region. The predicted cleavage site is at 55/54 aa. WaxS/WSlike is only 341/333 aa. We included the whole protein for both promoter plus target construct (Figure 9). Figure 11 shows WaxS is not localized in chloroplast. WSlike might be in chloroplast, but not only within chloroplast.

Materials and Methods

Plant Material

Plants (*Arabidopsis thaliana* ecotype Columbia) were grown in Sun Gro soil in 3" SQ pots and flats in a plant growth room under constant fluorescence white light at 22^{o}C.

In situ Hybridization

Arabidopsis siliques were harvested at 1, 3, 5, 7, 12 DAF, cut into 4 mm pieces, fixed, embedded and sectioned as previously described (Ke et al., 1997). DIG-labeled RNA probe (antisense and sense) were transcribed from the following insert + vector combinations: 676 nucleotides at 3'-end (positions 451 to 1114) of *CAC1A* cDNA in pBluescript SK (+/-), 789 nucleotides at 3'-end (positions 451 to 1217) of *CAC1B* cDNA in pSPORT 1, and 644 nucleotides at 5'end (positions -52 (from 5'UTR, not including intron in 5'UTR) to 592 (not including intron)) (GIBCO, Rockville, Maryland). After Proteinase K digestion, hybridization and washing at 65^{o}C, slides were treated with RNaseA to remove the RNA probe that had not hybridized with mRNA. The Boehringer Mannheim DIG nucleic acid detection kit (Boehringer Mannheim, Indianapolis, Indiana) was used for immunological detection (Canas et al., 1994). About 30 silique pieces were used in total. For each developmental stage, about 8 ovules per silique piece were examined, and the entire experiment was repeated three times.

Construction of Promoter/Promoter+Target::GFP/GUS Fusion Vectors

Promoter::GFP/GUS fusion constructs were made for all genes by cloning amplified promoter region into a binary vector. In cases where the intergenic regions are short (no more

than 1 kb), several constructs spanning various regions were made. The intergenic region for *CAC1A* is 884 bp. Three constructs were made for *CAC1A*: CAC1Ap1 (1124 bp upstream of ATG codon, obtained from Dr. Nikolau lab), CAC1Ap2 (884 bp upstream of ATG codon), and CAC1Ap3 (701 bp upstream of ATG codon). Primers used were CAC1Ap1F (5'-AAGCTTGATTGTAACCAAGACA-3'), CAC1Ap2F (5'-GTCTCTCTCTATATCACTCATTATG-3'), CAC1Ap3F (5'-GTTCATAACCAAGTAAATATGC-3'), and a common reverse primer: CAC1ApR (5'-TTCGTCTTCTTATTGTTATTGG-3'). The intergenic region for *CAC1B* is 818 bp. Two constructs were made for *CAC1B*: CAC1Bp1 (1244 bp upstream of -13 from ATG codon, obtained from Dr. Nikolau lab) and CAC1Bp2 (807 bp upstream of ATG codon). Primers used were CAC1Bp1F (5'-CAGGGTCATATGGAGCAAGC-3'), CAC1Bp1R (5'-GACGATGAAACCGAGGAAGT-3'); CAC1Bp2F (5'-CTCCGTTAAGTTAAGCCAAGA-3'), CAC1Bp2R (5'-ATTGTTGAGACAGTGGACGATG-3'). The intergenic region for *CACT* is 1353 bp. One construct was made for *CACT*: CACTp (1340 bp upstream of ATG codon). Primers used were CACTptarF (5'-TTGAGAAAAGAACTTATTGAC-3'), CACTpR (5'-TTCTAAGAAACAATAAATCAATC-3'). The intergenic region for *WaxS* is 280 bp. Two constructs were made for wax synthase: WaxSp1 (1551 bp upstream of ATG codon) and WaxSp2 (280 bp upstream of ATG codon). Primers used were WaxSp1tarF (5'-ACCTTGTTCCTCCGCCACT-3'), WaxSp2F (5'-ATCAGCACTCATTATCCTT-3'), and a common reverse primer: WaxSpR (5'-TTCTCAGATCTGTCGTTTGCTAA-3'). The intergenic region for *WSlike* is 1369 bp. One construct was made for wax synthase like: WSlikep: 1330 bp upstream of ATG codon. Primers used were WSlikeptarF (5'-TGATGATTTGGGAAGAGAACTA-3'), WSlikepR (5'-CTCTCAGATCTTTGTTTGTGTTG-3').

Promoter+target::GFP/GUS fusion constructs were made for *CACT*, *WaxS* and *WSlike* genes, one construct per gene. The promoter+target construct is the longest promoter region from the promoter constructs plus the whole protein (stop codon is not included): CACTp plus 1030 bp (including a 129-bp intron; including the whole 300 aa from CACT) for *CACT*; WaxSp1 plus 1024 bp (including the whole 341 aa from WaxS) for *WaxS*; WSlikep plus 1000 bp (including the whole 332 aa from WSlike) for *WSlike*. Primers used

were CACTptarF (5'-TTGAGAAAAGAACTTATTGAC-3'), CACTtarR (5'-AATCCCAAGCTTGACCTT-3'); WaxSp1tarF (5'-ACCTTGTTCCTCCGCCACT-3'), WaxStarR (5'-AATGAAGAAGTGAATAACTTGG-3'); WSlikeptarF (5'-TGATGATTTGGGAAGAGAACTA-3'), WSliketarR (5'-AATCGCTTAATGAACTCAACG-3').

Fragments including the promoter region or promoter plus target region were amplified by PCR from total DNA using the attB-adapted primers (5'-GGGGACAAGTTTGTACAAAAAAGCAGGCTTC-forward primer, 5'-GGGGACCACTTTGTACAAGAAAGCTGGGTC-reverse primer). After PCR-product purification, the amplified promoter or promoter plus target region was cloned first into an entry vector pDONR221, then into a binary vector pBGWFS7 (Karimi et al., 2002; Flanders Interuniversity Institute for Biotechnology, 2003) upstream of promoterless GFP/GUS gene (Gateway™ Cloning Technology, Invitrogen, CA, 2003). Gateway vectors were transformed into *Agrobacterium tumefaciens* strain GV3101.

Arabidopsis Transformation and Selection

Arabidopsis plants were transformed using the floral dip method (Clough and Bent, 1998). Transformed Arabidopsis plants were selected based on Bar resistance conferred by the T-DNA. T2 seeds from at least 15 independently-transformed lines for each construct were harvested for GUS and GFP screening.

Histochemistry and Microscopy

Transgenic T2 seedlings were germinated in soil in pots. For transgenic lines containing promoter::GFP/GUS constructs, plants were harvested at various stages of development. A minimum of three plants were harvested for each independently-transformed line; plants or organs from the same line were stained in the same tube. Organs were stained in GUS-staining solution (Triton/Ethanol stock (Triton X-100: Ethanol: Water, 1:4:5): 0.5M PH7.0 KPO4 Buffer: 10mg/ml X-gluc in dimethyl sulfoxide: 0.1M PH7.0 Potassium ferricyanide: 0.1M PH7.0 Potassium ferrocyanide, 5:470:25:2:2) at room temperature overnight or for three hours (Jefferson, 1987**)**. Staining patterns were observed and

documented using an Olympus stereomicroscope in the Bessey Microscopy Facility (Ames, IA).

GFP Fluorescence Microscopy

For promoter-target construct, young leaves (leaves three to six) from 15 DAI (days after planting) plants were placed in water on slides, covered with cover slips. Slides were observed and documented under a Leica TCS NT laser scanning microscope system in the Confocal Microscopy Facility (Ames, IA). Under confocal microscope, GFP signal shows green color and the auto fluorescence shows red color. The Argon/Krypton laser and 488 nm/568 nm FITC/TRITC wavelengths were used (Omnichrome, Chino, CA). We found that yellowish or damaged or dying part of the leaf would show green color under confocal microscope. Only the healthy, green leaves were selected to detect real GFP signal. Microscope digital images were processed in Adobe Photoshop 7.0 (Adobe, San Jose, CA).

LITERATURE CITED

Alban C, Baldet P, Douce R (1994) Localization and characterization of two structurally different forms of acetyl-CoA carboxylase in young pea leaves, of which one is sensitive to aryloxyphenoxyproprionate herbicides. Biochem J **300:** 557–565

Canas LA, Busscher M, Angenent GC, Beltran J-P, van Tunen AJ (1994) Nuclear localization of the petunia MADS box protein FBP1. The Plant Journal **6:** 597–604

Choi J-K, Yu F, Wurtele ES, Nikolau BJ (1995) Molecular cloning and characterization of the cDNA coding for the biotin-containing subunit of the chloroplastic acetyl-coenzyme A carboxylase. Plant Physiol **109:** 619–625

Fatland BL, Nikolau BJ, Wurtele ES (2005) Reverse genetic characterization of cytosolic acetyl-CoA generation by ATP-citrate lyase in Arabidopsis. Plant Cell **17:** 182–203

Jefferson RA (1987) Assaying chimeric genes in plants: The GUS gene fusion system. Plant Mol Biol Rep **5:** 387–405

Kahn AA, Kolattukudy PE (1973) Control of synthesis and distribution of acyl moieties in etiolated *Euglena gracili*s. Biochemistry **12:** 1939–1948

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

Ke J, Choi JK, Smith M, Horner HT, Nikolau BJ, Wurtele ES (1997) Structure of the CAC1 gene and in situ characterization of its expression. The Arabidopsis thaliana gene coding for the biotin- containing subunit of the plastidic acetyl-coenzyme A carboxylase. Plant Physiol **113:** 357–365

Ke J, Wen TN, Nikolau BJ, Wurtele ES (2000) coordinate regulation of the nuclear and plastidic genes coding for the subunits of the heteromeric acetyl-coenzyme A carboxylase. Plant Physiol **122:** 1057–1071

Kolattukudy PE, Rogers L (1978) Biosynthesis of fatty alcohols, alkane-1,2-diols and wax esters in particulate preparations from the uropygial glands of white-crowned sparrows *(Zonotrichia Leucophyrus*). Arch Biochem Biophys **191:** 244–258

Kolattukudy PE, Rogers L (1986) Acyl-CoA reductase and acyl-CoA: fatty acyl transferase in the microsomal preparation from the bovine meibomian gland. J Lipid Res **27:** 404–411

Konishi T, Sasaki Y (1994) Compartmentalization of two forms of acetyl-CoA carboxylase in plants and the origin of their tolerance toward herbicides. Proc Natl Acad Sci USA **91:** 3598–3601

Lardizabal KD, Metz JG, Sakamoto T, Hutton WC, Pollard MR, Lassner MW (2000) Purification of a jojoba embryo wax synthase, cloning of its cdna, and production of high levels of wax in seeds of transgenic Arabidopsis. Plant Physiol **122:** 645–655

Lassner MW, Lardizabal K, Metz JG (1996) A jojoba β-ketoacyl-CoA synthase cDNA complements the canola fatty acid elongation mutation in transgenic plants. Plant Cell **8:** 281–292

Lawand S, Dorne AJ, Long D, Coupland G, Mache R, Carol P (2002) Arabidopsis A BOUT DE SOUFFLE, which is homologous with mammalian carnitine acyl carrier, is required for postembryonic growth in the light. Plant Cell **14:** 2161–2173

Metz JG, Pollard MR, Anderson L, Hayes TR, Lassner MW (2000) Purification of a jojoba embryo fatty acyl-coenzyme A reductase and expression of its cDNA in high erucic acid rapeseed. Plant Physiol **122:** 635–644

Qian HR (2002) Molecular genetic studies of acetyl-CoA carboxylase and 3-methylcrotonyl-CoA carboxylase in plants. Ph.D. dissertation. Iowa State University, Ames, IA

Thelen JJ, Mekhedov S, Ohlrogge JB (2001) Brassicaceae express multiple isoforms of biotin carboxyl carrier protein in a tissue-specific manner. Plant Physiol **125:** 2016–2028

Wu XY, Moreau RA, Stumpf PK (1981) Studies of biosynthesis of waxes by developing Jojoba seed. III. Biosynthesis of wax esters from acyl-CoA and long chain alcohols. Lipids **6:** 897–902

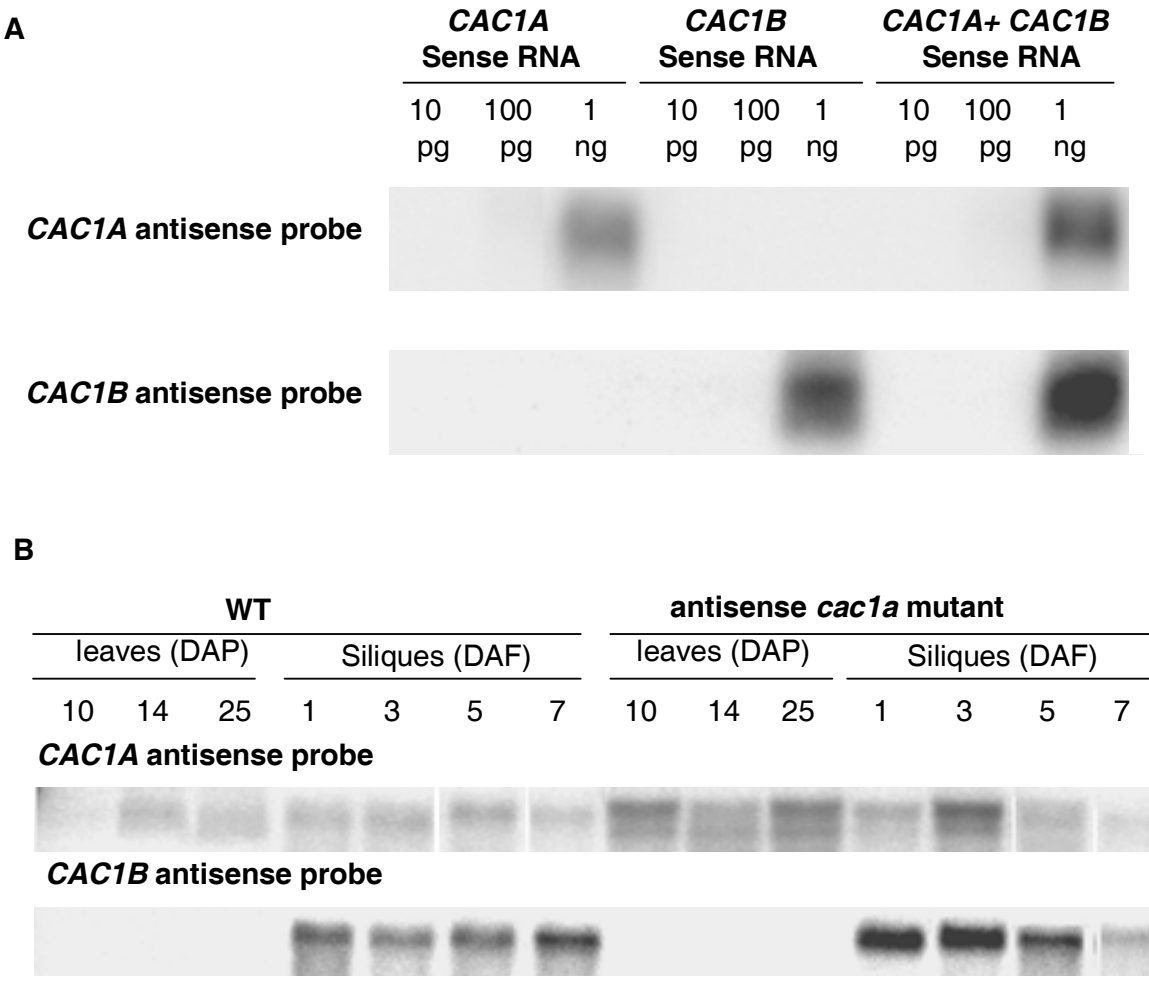


Figure 1. *CAC1A* and *CAC1B* probes do not cross hybridize; *CAC1A* and *CAC1B* RNAs accumulate differently in leaves.
(A) Northern blot analyses for *CAC1A* and *CAC1B* sense RNA.
(B) Northern blot analyses for *CAC1A* and *CAC1B* transcripts in leaves and siliques of WT and antisense *cac1a* mutant. Only *CAC1A* RNA is accumulated in both leaves and siliques. *CAC1B RNA* is accumulated in siliques. There is still *CAC1A* RNA accumulation in antisense *cac1a* mutant. The level of *CAC1B* RNA accumulation is not affected in the antisense *cac1a* mutant.

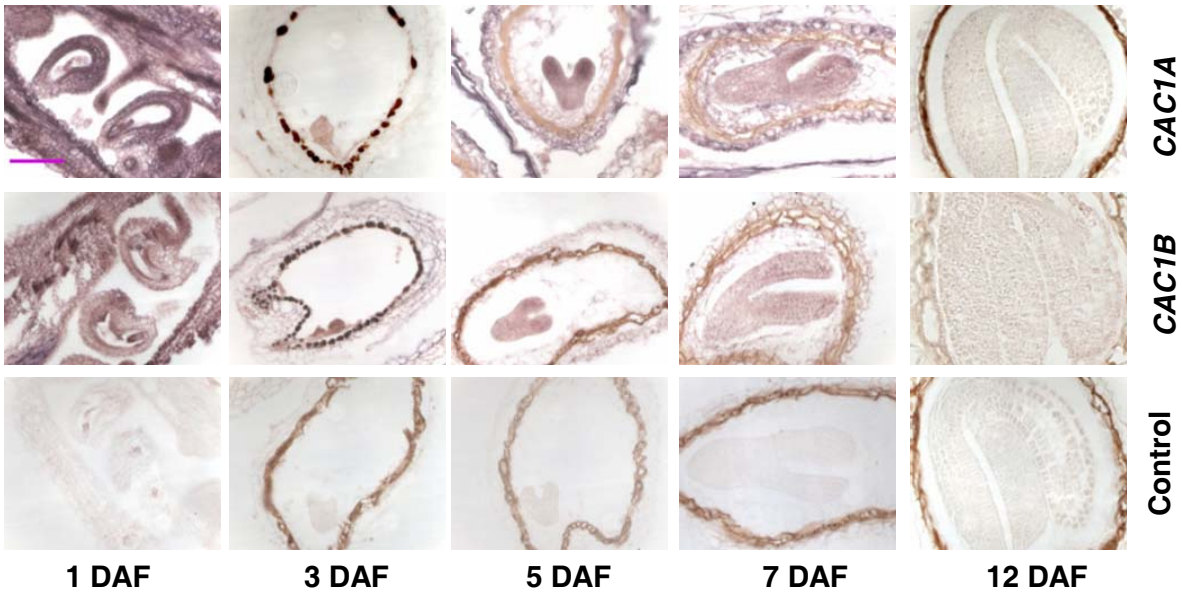


Figure 2. *CAC1A* and *CAC1B* RNAs accumulate with similar spatial and t uyemporal patterns during embryo development.
In situ location of the *CAC1A* and *CAC1B* mRNAs during embryo development (Bar=50µm). In rapidly enlarging siliques (1DAF), the sporophyte within the ovule is a zygote or several-celled embryo. *CAC1A* and *CAC1B* mRNAs are evenly distributed throughout the silique, including the silique walls and developing ovules. By 3 DAF, siliques have ceased enlargement. Their mRNAs are concentrated in the rapidly growing globular embryos, and are dramatically reduced in silique walls and ovule integuments. At 5 to 7 DAF, the time of maximal seed oil accumulation, their mRNAs are highly concentrated in the embryos. By 12 DAF, embryos have ceased growing and are approaching desiccation; at this stage, *CAC1A* and *CAC1B* mRNAs are no longer detected in the embryos or siliques.

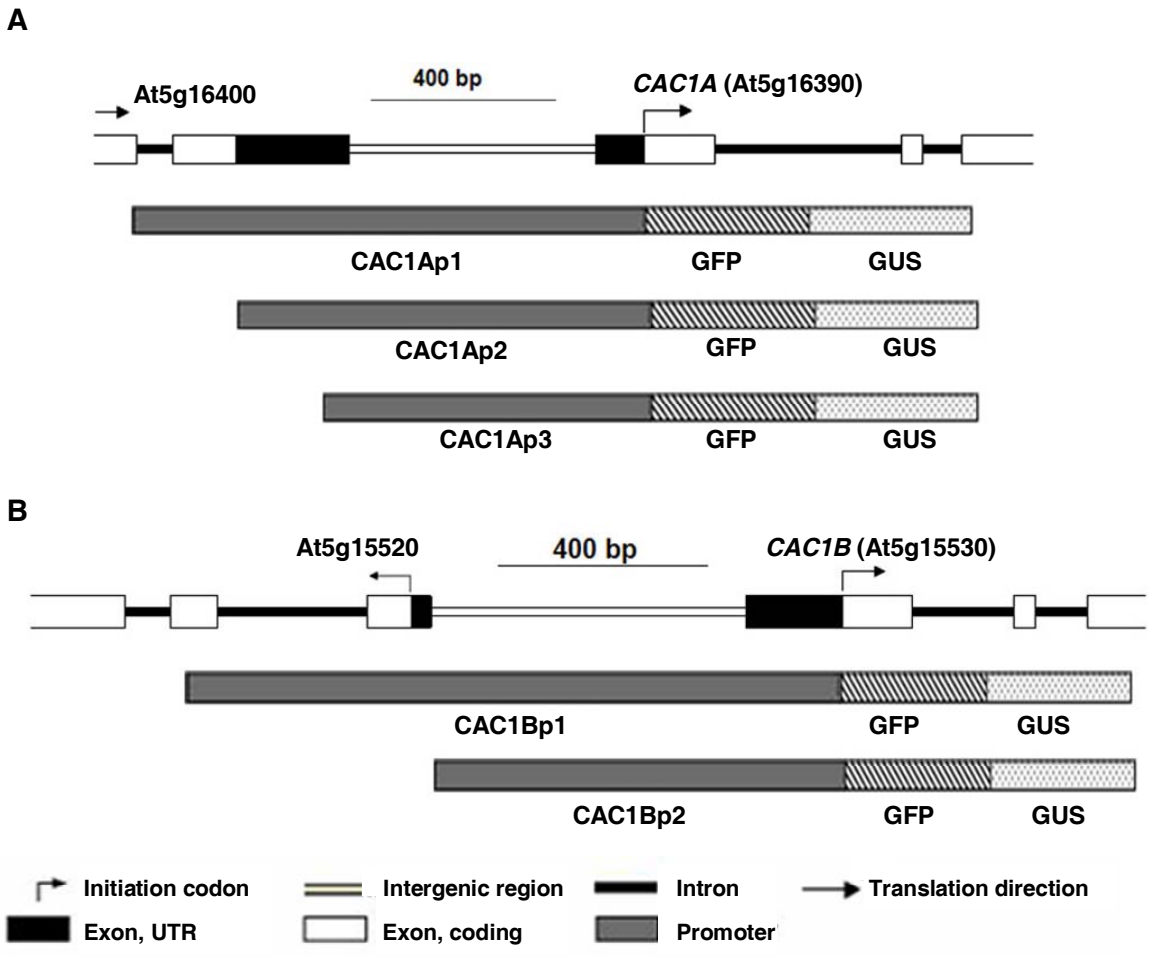


Figure 3. Promoter::GFP/GUS constructs for *CAC1A* and *CAC1B*.
(A) Three constructs for *CAC1A*: CAC1Ap1 (1124 bp upstream of ATG codon, obtained from Dr. Nikolau lab), CAC1Ap2 (884 bp upstream of ATG codon), and CAC1Ap3 (701 bp upstream of ATG codon).
(B) Two constructs for *CAC1B*: CAC1Bp1 (1244 bp upstream of -13 from ATG codon, obtained from Dr. Nikolau lab) and CAC1Bp2 (807 bp upstream of ATG codon).

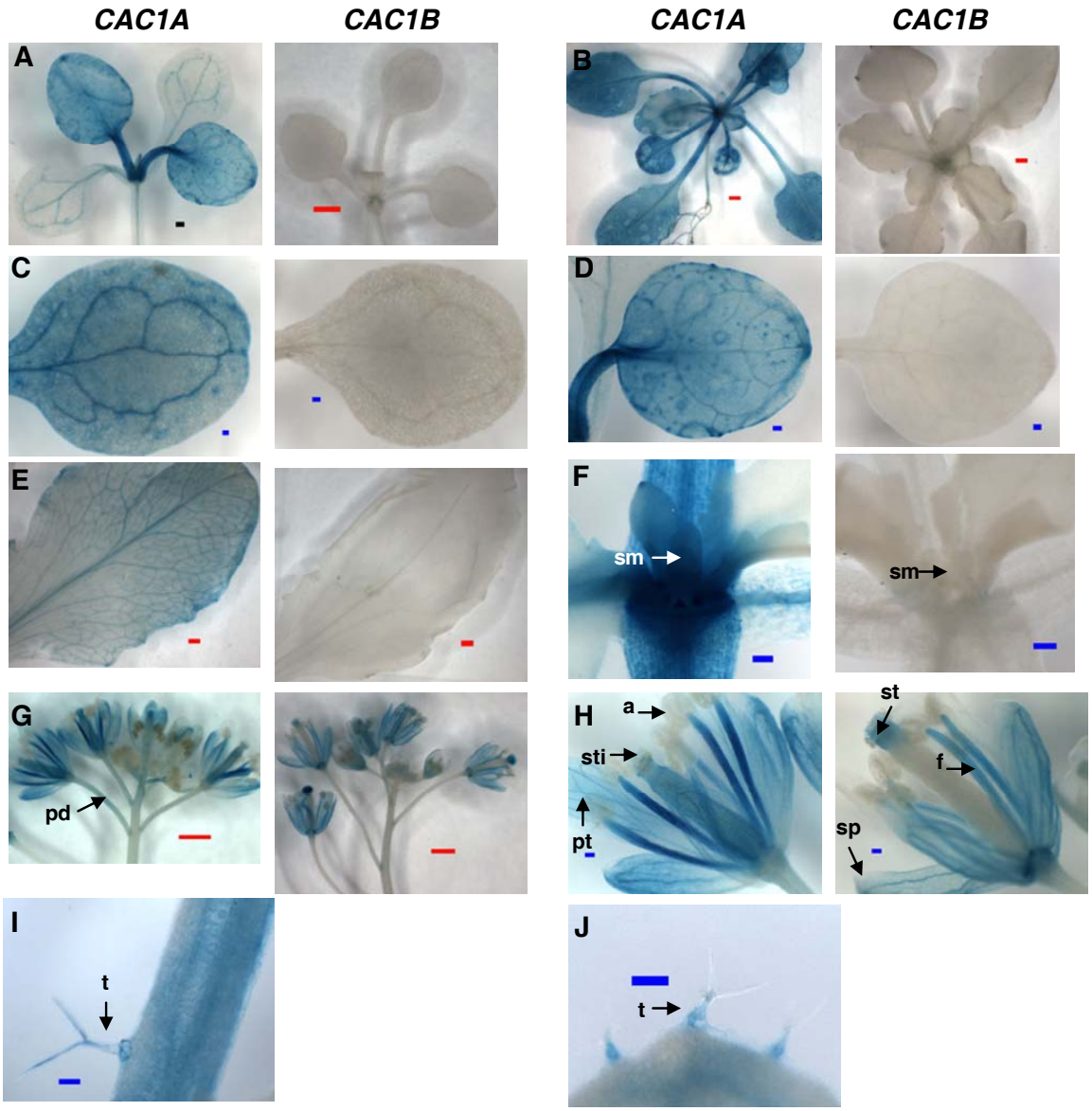


Figure 4.

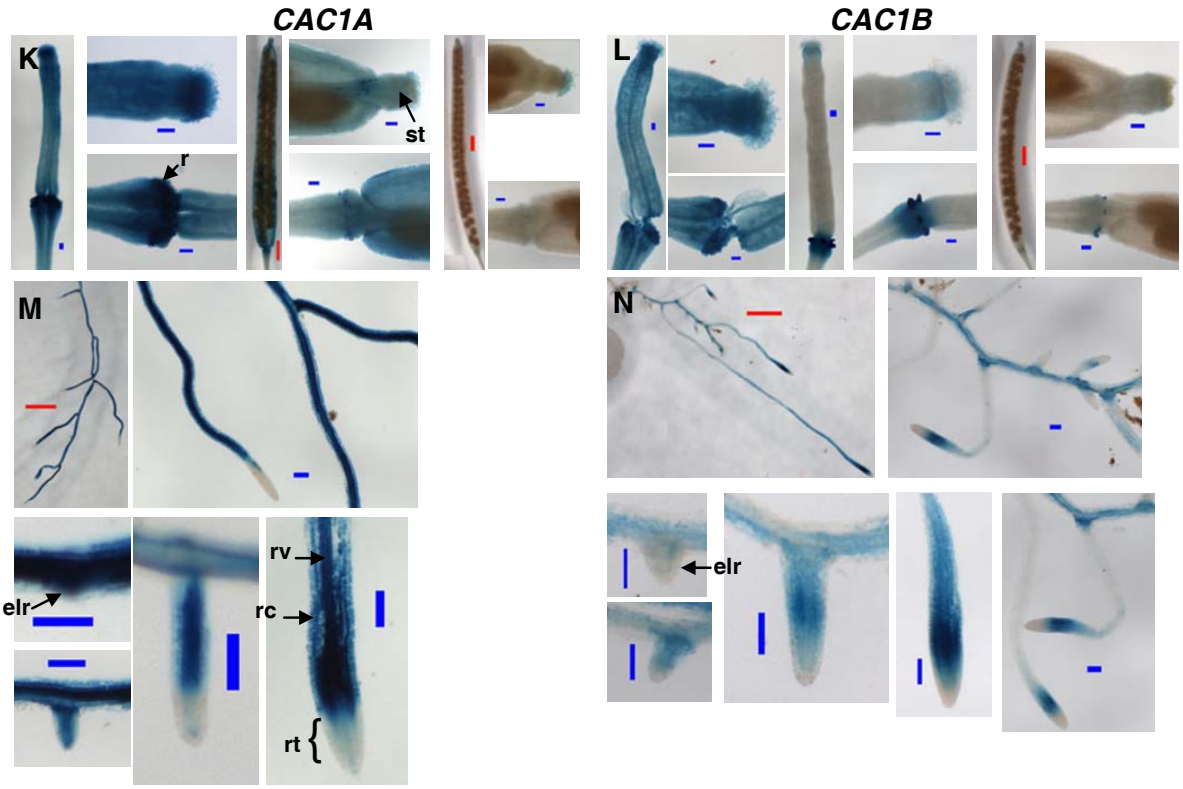


Figure 4. (continued). *CAC1A* and *CAC1B* are differently expressed. *CAC1A* is expressed more widely than *CAC1B*. Where there is *CAC1B* expression, there is higher expression of *CAC1A*.

CAC1A and *CAC1B* expression in young seedling (A), plant (B), cotyledon (C), young leaf (D), mature leaf (E), shoot meristem (F), flowers and buds (G), flower (H), stem trichome (I), leaf trichome (J), siliques (K, L), and roots (M, N). *CAC1A* is expressed in cotyledons, leaves (leaf vasculature and leaf mesophyll cell) and shoot meristem from young seedling stage to post-bolting. It is in flowers in sepal, petal, filament, stigma, young silique, young embryo, receptacle, and pedicle. It is in stem and trichomes. It is not in anther, pollen and not in old silique. It is highly expressed in roots except the root tip. *CAC1B* is not expressed in seedling, not in cotyledons or leaves, not in shoot meristem. It is in flowers in sepal, filament, stigma, young silique, young embryo and receptacle. It has similar expression pattern in root as *CAC1A* but with reduced density. Where there is *CAC1B* expression, there is higher expression of *CAC1A*. The expression in silique and embryo is consistent with *in situ* result. The GUS expression pattern also supports the hypothesis that BCCP1 is important for htACCase activity *in planta*, but BCCP2 is dispensable. a=anther, elr=emerging lateral root, f=filament, pd=pedicle, pt=petal, r=receptacle, rc=root cortex, rt=root tip, rv=root vasculature, sm=shoot meristem, sp=sepal, sti=stigma, st=style, t=trichome. Red bar=1mm, black bar=0.2mm, blue bar=0.1mm.

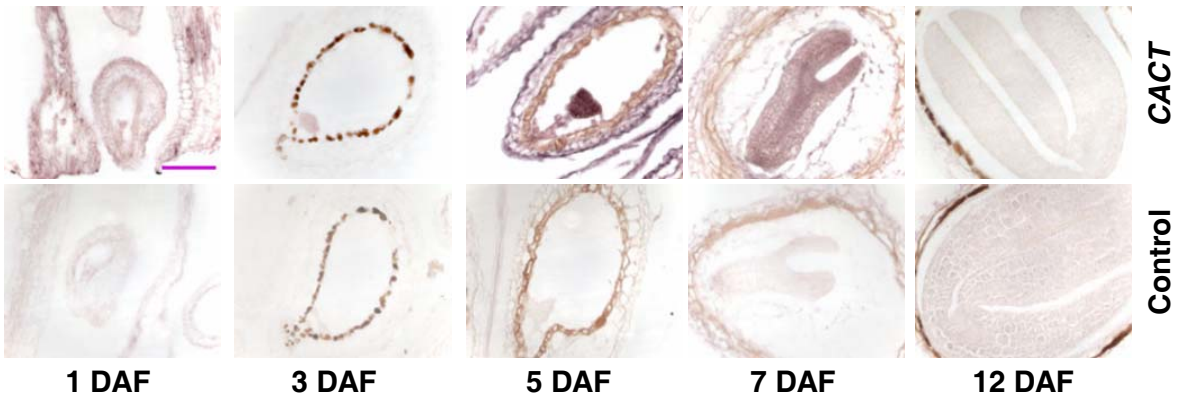


Figure 5. The pattern of *CACT* expression is consistent with a role in embryogenesis. *In situ* location of the *CACT* mRNAs during embryo development (Bar=50μm). *CACT* is expressed in *Arabidopsis* siliques. *CACT* mRNA is distributed all over the silique tissues at 1 DAF. It is reduced in the silique walls and mainly concentrated in the embryo at 3 DAF. From 5 DAF to 7 DAF, it is highly concentrated in the embryo, reaching peak at 5 DAF. It decreases after 7 DAF and is almost non-detectable at 12 DAF.

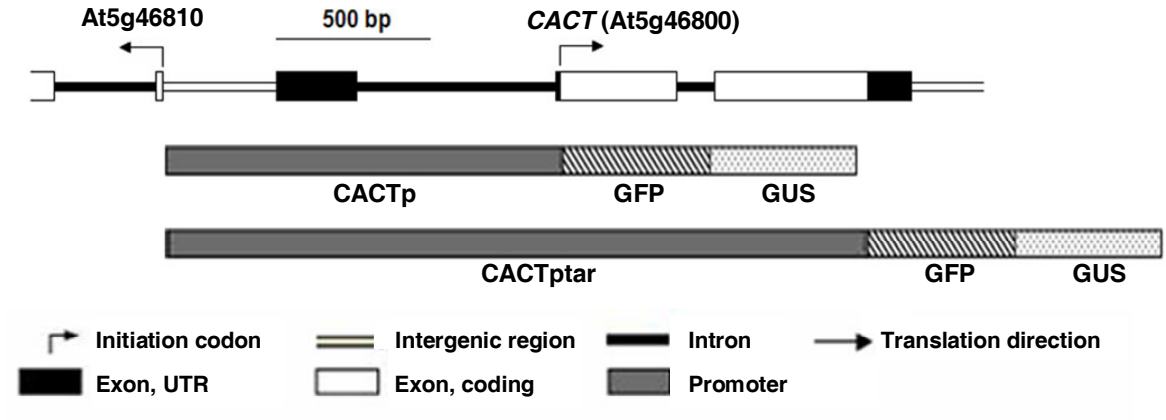


Figure 6. *CACT* promoter/promoter+target::GFP/GUS constructs.
Promoter::GFP/GUS construct CACTp: 1340 bp upstream of ATG codon.
Promoter+target::GFP/GUS construct CACTptar: promoter region + whole protein (not include the stop codon).

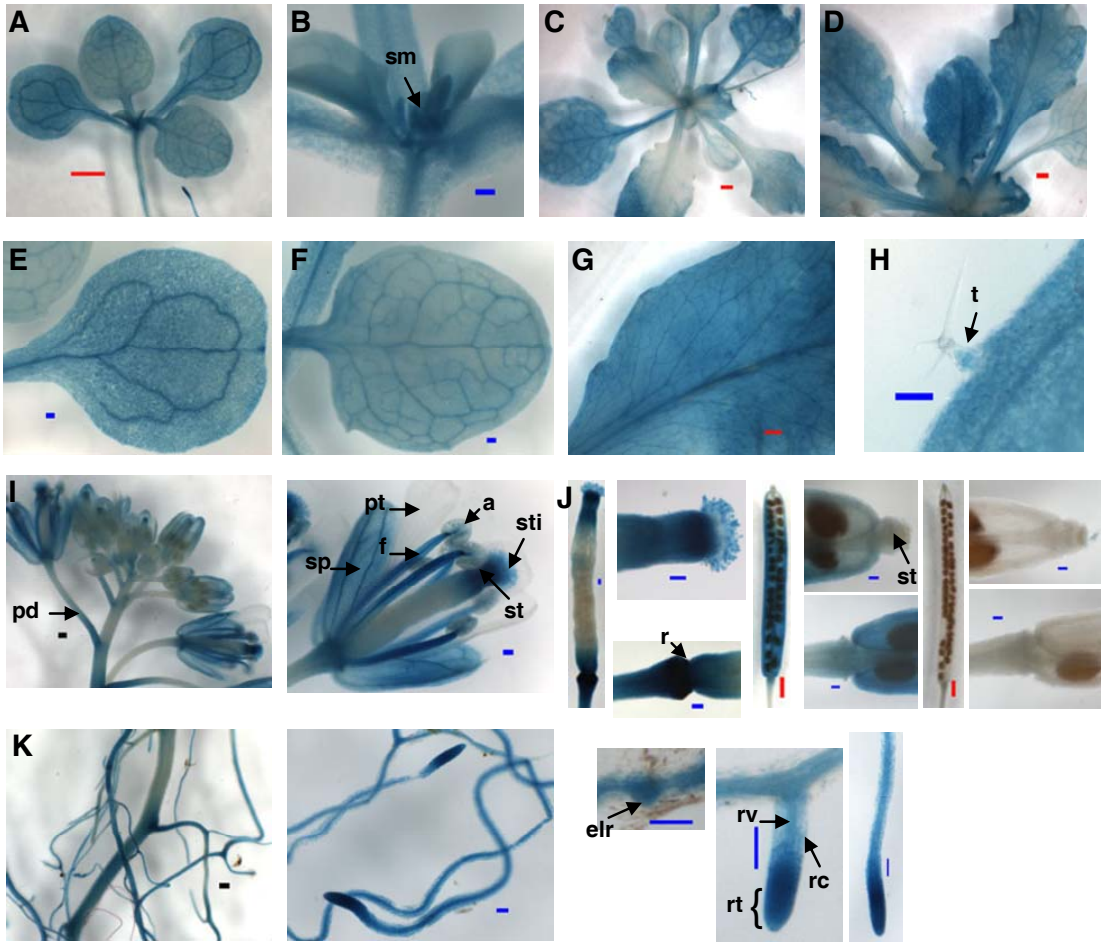


Figure 7. *CACT* expression is similar to that of *CAC1A*, which is consistent with a role in embryogenesis. *CACT* expression in seedling (A), shoot meristem (B), plant before bolting (C), rosette after bolting (D), cotyledon (E), young leaf (F), mature leaf (G), trichome at leaf edge (H), flowers and buds (I), siliques (J), roots (K). *CACT* is expressed in cotyledons, leaves (leaf vasculature and leaf mesophyll cell) and shoot meristem from young seedling to mature plant. It is in flowers in sepal, filament, stigma, young silique, young embryo, receptacle, and pedicle. It is in stem and trichomes. It is not in anther, and not in old silique. It is highly expressed in roots. What makes it differ from *CAC1A* is that it is expressed in pollen but not in petal. It is expressed in root tip. a=anther, elr=emerging lateral root, f=filament, pd=pedicle, pt=petal, r=receptacle, rc=root cortex, rt=root tip, rv=root vasculature, sm=shoot meristem, sp=sepal, sti=stigma, st=style, t=trichome. Red bar=1mm, black bar=0.2mm, blue bar=0.1mm.

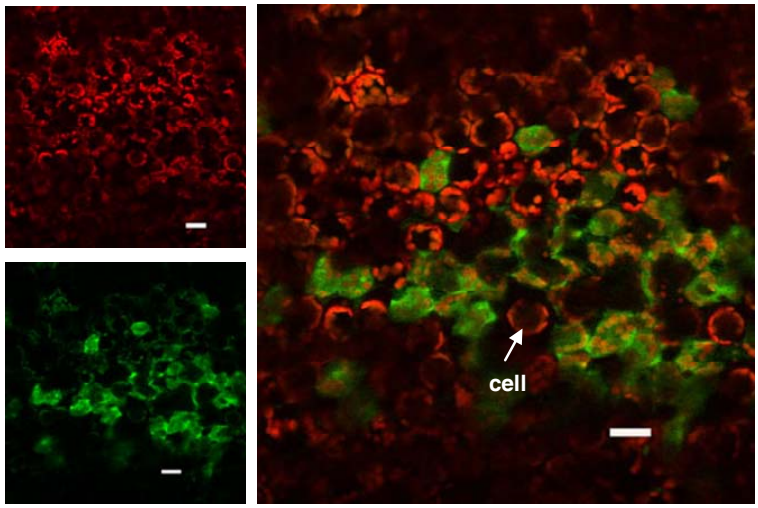


Figure 8. CACT is not localized in plastids.
Left panel: top, auto-fluorescence (red); bottom: GFP signal (green). Right panel, *CACT* promoter-transcribed sequence-GFP fusion in leaf. White bar=20μm. The method limits us from further identification of its sub-cellular localization.

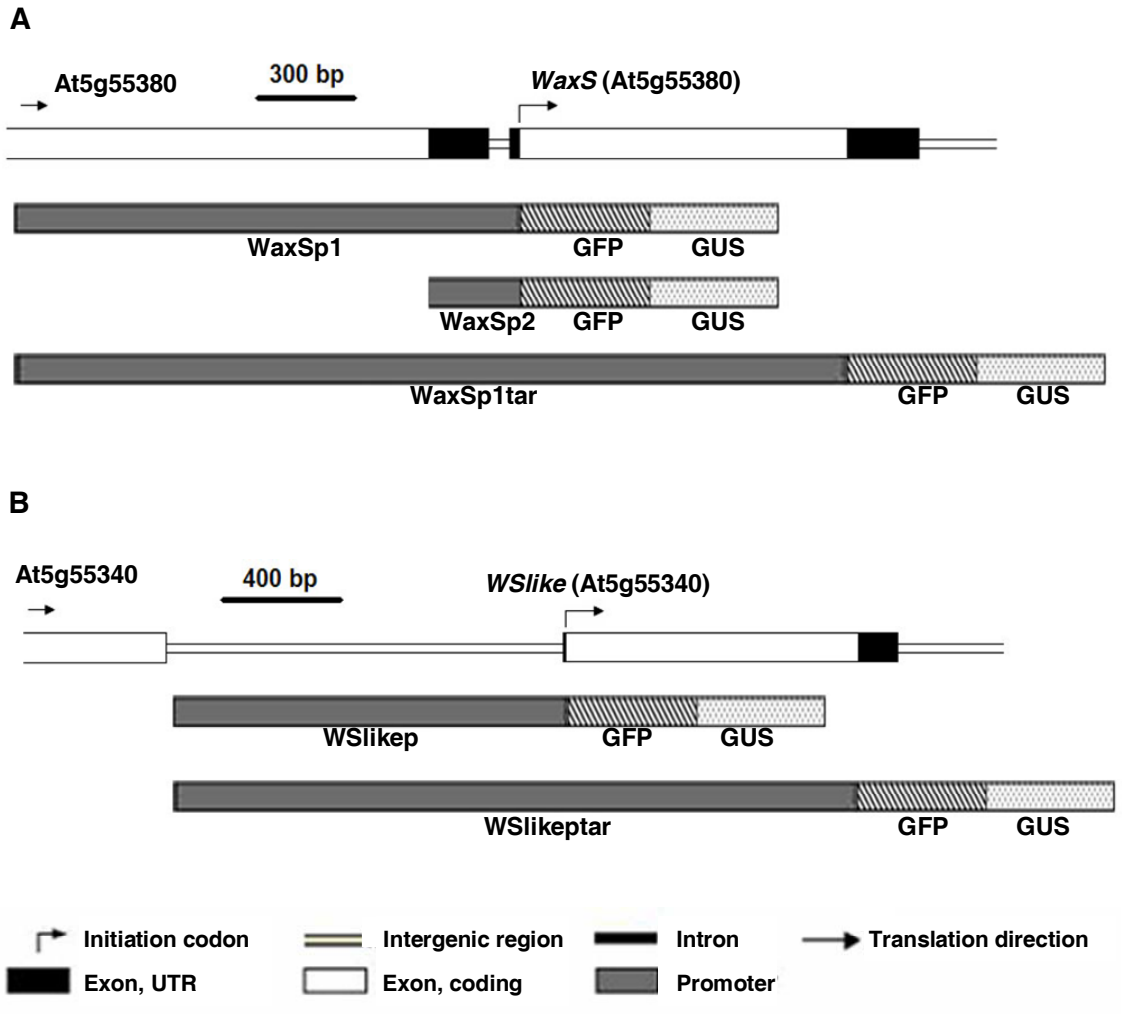


Figure 9. Wax synthase and wax synthase like genes promoter/promoter + target::GFP/GUS constructs.
(A) Constructs for wax synthase: WaxSp1: 1551 bp upstream of ATG codon; WaxSp2: 280 bp upstream of ATG codon. WaxSp1tar: the promoter region plus whole protein (not include the stop codon).
(B) Constructs for wax synthase like: WSlikep: 1330 bp upstream of ATG codon; WSlikeptar: promoter region plus whole protein (not include the stop codon).

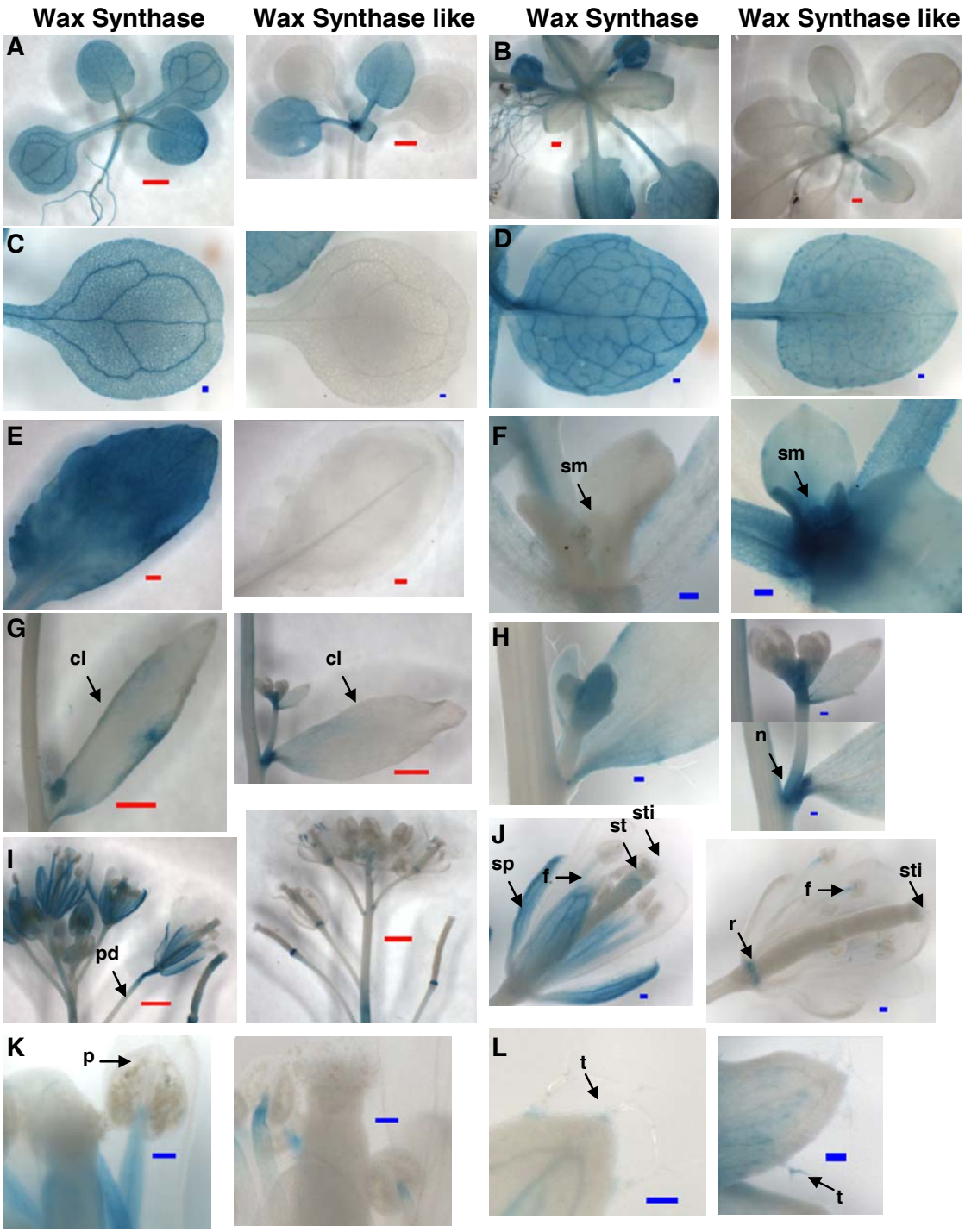

Wax Synthase
Wax Synthase like
Wax Synthase
Wax Synthase like
A
B
C
D
E
F
sm
sm
G
cl
cl
H
n
I
pd
J
sp
f
st
sti
f
sti
r
K
p
L
t
t

Figure 10.

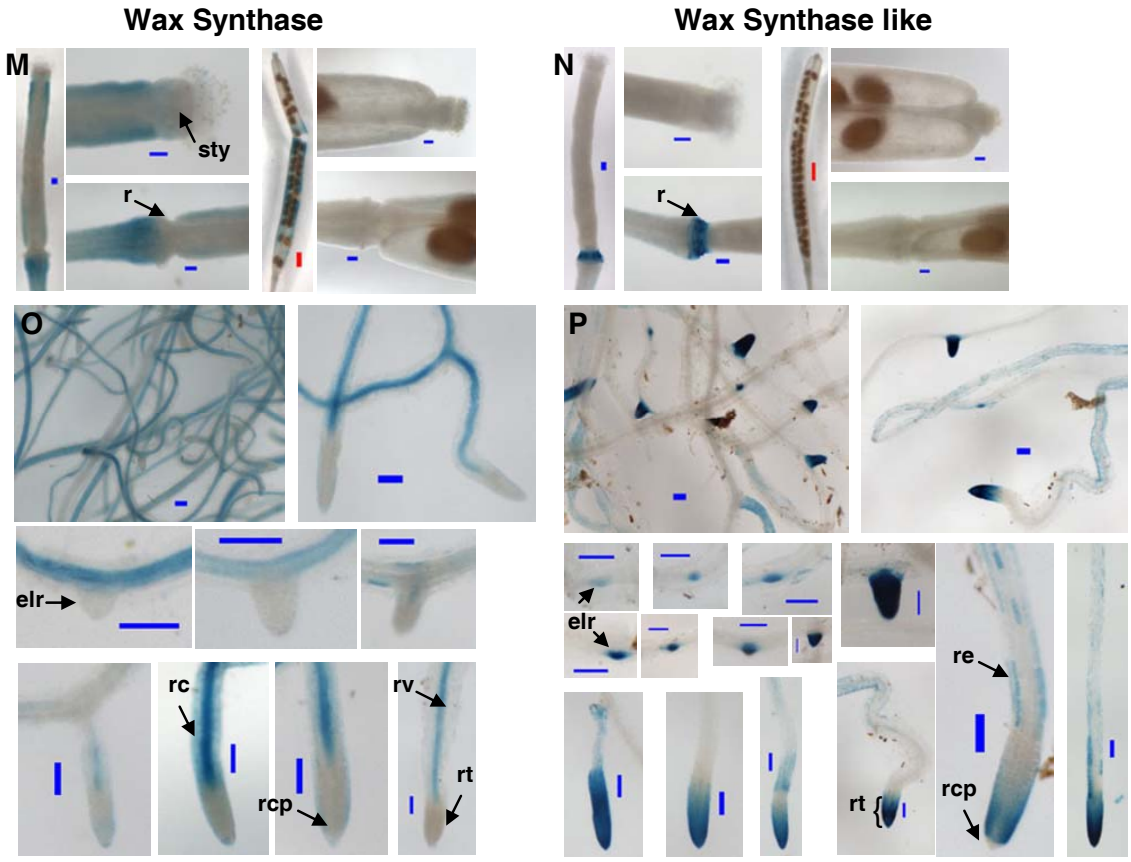
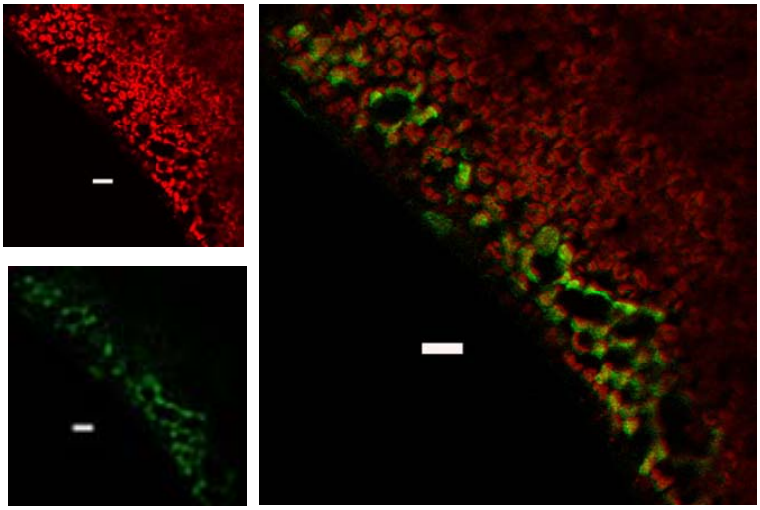


Figure 10. (continued).

Figure 10. (continued). *WaxS* (At5g55340) and *WSlike* (At5g55340) have different temporal and spatial expression patterns. The expression of *WaxS* and *WSlike* is complementary. Expression in seedling (A), plant before bolting (B), cotyledon (C), young leaf (D), mature leaf (E), shoot meristem (F), cauline leaf and node (G, H), flowers and buds (I and J), pollen (K), trichome at lef edge (L), siliques (M and N), and roots (O and P). *WaxS* has strong expression in cotyledons, and true leaves. In leaf mesophyll cells and leaf vasculature. It is almost in all vasculatures. It is in shoot base and petiole. It is not in shoot meristem. *WaxS* is expressed strongly in primary root, but not in germinating root. It is not expressed in root tip, but is at root cap. It is expressed in root vasculature and less in cortex. At just before bolting, it shows strongest in older leaves, but much less in young expanding leaves, it is not in shoot bolting buds. After plants bolting, *WaxS* has low expression in stems and strong expression in older leaves. It has strong expression in opened flower and young buds in sepals, pedicles (opened flowers), and also inside the flower (filament except the top part, and mature pollen), less in stalks and stigma. It is not in petal, not in anther, nor in receptacle. It is also in young and old silique wall. It is in cauline leaf trichome base. *WSlike* is not in cotyledon. It is expressed in shoot meristem and young leaves. At just before bolting, it is expressed in shoot meristem and young leaves. It is in leaf mesophyll cells, not obvious in veins. It is in bolting buds. After bolting, it has very weak staining in shoot meristem and young leaves. It is concentrated at stems just under flowers, cauline leaf and stems, in meristem around the cauline leaf. It is in receptacle and top filament. It is not in stigma, nor pollen. It is in cauline leaf trichome top. It is not in silique. It is expressed extraordinary in root tip, in germinating roots. It is not all over the roots, it is on root epidermal cells, root nodes surface. It is in secondary roots. The relationship of the expressions of *WaxS* and *WSlike* is different from the expressions of *CAC1A* and *CAC1B*. Unlike *CAC1A* expression overlaps the expression of *CAC1B*, the expression of *WaxS* and *WSlike* is complementary. Expression of *WSlike* is kind of opposite to *WaxS*. *WSlike* is usually expressed in where there is no *WaxS* expression. a=anther, cl=cauline leaf, elr=emerging lateral root, f=filament, n=node, p=pollen, pd=pedicle, pt=petal, r=receptacle, rc=root cortex, rcp=root cap, re=root epidermal, rt=root tip, rv=root vasculature, sm=shoot meristem, sp=sepal, sti=stigma, st=style, t=trichome. Red bar=1mm, black bar=0.2mm, blue bar=0.1mm.

Wax Synthase



Wax Synthase like

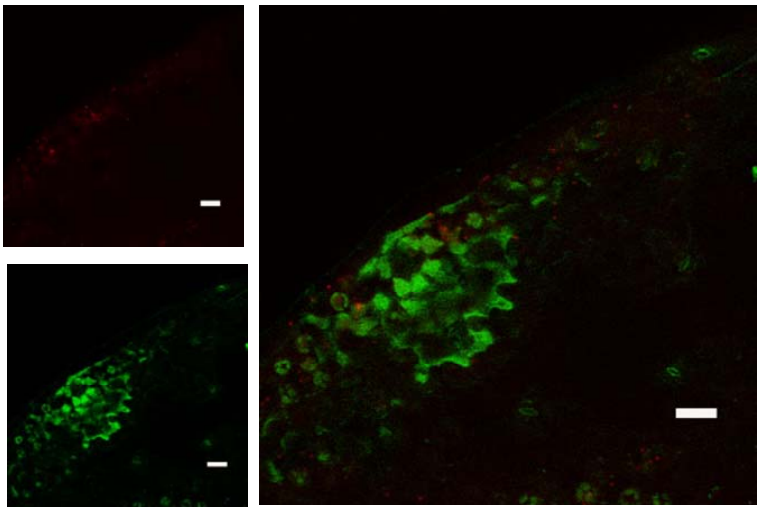


Figure 11. WaxS protein does not appear to be in chloroplast. WSlike protein might be in the chloroplast, but not exclusively in the chloroplast.
Left panel: top, auto-fluorescence (red); bottom: GFP signal (green). Right panel, promoter-transcribed sequence-GFP fusion in leaf. White bar=20μm. The instrument limits us from further identification of its sub-cellular localization.

Table 1. The promoter::GUS expression of genes related to acetyl-CoA/biotin network.

Tissue		***CAC1A***	***CAC1B***	***CACT***	***WaxS***	***WSlike***
Root	Tip	-	-	+	+	-
	Vasculature	+	Less	+	+	-
	Epidermis/Cortex	+	Less	+	+	-
	Root surface	+	-	+	-	+
	Root meristem	-	-	+	-	+
Cotyledon	Petiole	+	-	+	+	-
	Vasculature	+	-	+	+	-
	Mesophyll	+	-	+	+	-
	Hydathode	+	-	+	+	-
Leaf	Shoot meristem	+	-	+	-	+
	Petiole	+	-	+	+	+
	Vasculature	+	-	+	+	Less
	Mesophyll	+	-	+	+	Less
	Hydathode	+	-	+	+	+
	Trichome	+	-	+	+	+
Cauline leaf	Petiole	+	-	+	-	+
	Vasculature	+	-	+	Part	+
	Mesophyll	+	-	+	Part	+
	Hydathode	+	-	+	+	+
Stem	Vasculature	+	-	+	+	-
	Cortex	+	-	+	+	-
	Epidermis	+	-	+	-	-
Flower	Pedicle	+	+	+	+	+
	Receptacle	+	+	+	-	+
	Sepal	+	+	+	+	-
	Petal	+	+	-	-	-
	Filament	+	+	+	+(Base)	+(Top)
	Anther	-	-	+	-	-
	Pollen	-	-	+	+	-
	Ovary wall	+	+	+	+	-
	Style	+	+	+	+	-
	Stigma	+	+	+	+	-
Silique	Wall	+(Young)	+(Young)	+(Young)	+	-
	Seed/Embryo	+(Young)	+(Young)	+(Young)	-	-

"+" means there is signal. "-" means no signal is detected.

APPENDIX B. SOYBEAN SEED COMPOSITION

About 35 to 50% protein, 15 to 25% lipid and about 10% nonstructural carbohydrate are contained in typical mature soybean seeds. Soybean seeds provide about 25% of the world's supply of oils and 66% of vegetable proteins. The oil and protein accumulate in the cotyledons during seed development. Seed composition is a complex genetically determined trait that is also influenced by environmental modulators (Brummer et al. 1997). Once we understand the molecular basis for genetic and environmental variation in oil and protein composition, we can think about how to approach altering soybean seed composition for new food and industrial uses.

Near-isogenic lines that vary in seed composition can be used to identify genes whose expressions vary in correlation with seed composition. These genes are candidates for seed-composition quantitative trait loci (QTL) genes (Diers et al. 1992, Mansur et al. 1993, Brummer et al., 1997). We have a set of unique lines that are 94% isogenic to a common parent and vary by nearly 10% (from approximately 30 to 40%) in seed protein content. Evans is a commercial line relatively low in protein. PI438.472 is even lower in protein. PI153.296 is high in protein. The crosses between Evans and PIs are back-crossed three times to Evans. Compare Evans and BC3 progeny that differ in seed composition will identify seed-composition QTLs associated with either high or low seed protein content.

We also use another set of lines to identify soybean genes that maintain the stability of seed composition in response to different growth temperature. Evans is low in protein and increases seed protein at high temperature. Proto is a high protein line that also increases seed protein at high temperature. PI132.217 contains middle protein content and does not change protein content when temperature changes. Those genes that alter their expression in response to temperature in Evans and Proto but do not alter in PI132.217 may have a critical role in maintaining a stable seed-composition in response to different growth temperatures, may code for factors that regulate the temperature-induced changes in seed-composition.

Soybean pod growth precedes seed growth (Figure 1). Dr. Westgate Lab provided the growth, lipid and protein accumulation data (Figure 2). According to this data, we chose different stages of seeds for microarray analysis. I extracted total RNA and poly A^+ RNA

from the seeds and gave them to Dr. Periappuram (total RNA and poly A^+ RNA) and the ISU GeneChip facility (total RNA). Dr. Periappuram made the cDNA chips and did the hybridization. He provided the cDNA microarray data. GeneChip facility provided the Affymetrix soybean genome array data. The soybean seeds and RNA samples information is shown in Table 1.

I did northern blots to check the quality of total RNA and poly A^+ RNA. Figure 3 shows the sequences of northern probes of *ACLA* and *ACLB*. Figure 4 shows that *ACLB* mRNA is highly accumulated in early stage of seed development of Evans (around 25DAF), PI132.217 (around 30DAF) and Proto (around 25DAF). *ACLA* mRNA is highly accumulated at about the same stages of Evans, PI132.217 and Proto seed development. The high expression of *ACLA* and *ACLB* might be related to the synthesis of steroids or other metabolites derived from cytosolic acetyl-CoA in seeds.

Table 2 shows the information about the Northern probes. The mRNA profiling patterns are confirmed by northern blots (Figure 5 and 6).

Plant Material

Glycine max genetic determination of seed composition experiment - Evans is a commercial line of soybean relatively low in protein. PI438.472 is even lower in protein. PI153.296 is high in protein. The temperature was kept at 27/20^{o}C in the growth chamber.

Glycine max temperature modulation of seed composition experiment - Evans is a commercial line that increases seed protein at high temperature. Proto is a high protein line that also increases seed protein at high temperature. PI132.217 contains middle protein content and does not change protein content when temperature changes. PI132.217, Proto and Evans were grown under controlled environmental conditions at 27/20^{o}C until flowering and pod set were complete. Then one third of them were kept at 27/20^{o}C, one third were grown at 35/27^{o}C, and one third of them were grown at 20/12^{o}C until physiological maturity.

Miocroarray Design

Soybean plants were grown in pots, three plants/pot. The pots were placed randomly in the growth chamber. One growth chamber was one replication; three growth chambers

were used. Pods were collected from nodes 6-10 on the main stem. Uniform seeds were selected, frozen immediately in liquid N_2, and stored at -80^{o}C. Each developmental stage of each genotype had three replicates. Total RNA was extracted from soybean seeds by TRIzol RNA isolation method and poly A^+ RNA was isolated from total RNA with Qiagen Oligotex mini Prep kit. Northern blots were done to check the quality of the total RNA and poly A^+ RNA. cDNA probes containing Cy3- or Cy5-labeled nucleotides were synthesized and each individual DNA microarray chip was hybridized with pairs of probes (Cy3- or Cy5-labeled) by Dr. Periappuram. Northern blots were used to verify the profiling patterns of certain genes.

LITERATURE CITED

Brummer EC, Graef GL, Orf J, Wilcox JR, Shoemaker RC (1997) Mapping QTL for seed protein and oil content in eight soybean populations. Crop Science **37:** 370–378

Diers BW, Keim P, Fehr WR, Shoemaker RC (1992) RFLP analysis of soybean seed protein and oil content. Theor Appl Genet **83:** 608–612

Mansur LM, Lark KG, Kross H, Oliveira A (1993) Interval mapping of quantitative trait loci for reproductive, morphological, and seed traits of soybean (*Glycine max* L.). Theor Appl Genet **86:** 907–913

Figure 1. Soybean (PI132.217) pod and seed development.
After the pod almost finishes growth at around 15DAF, the seed begins its rapid growth.
Bar = 1 cm.

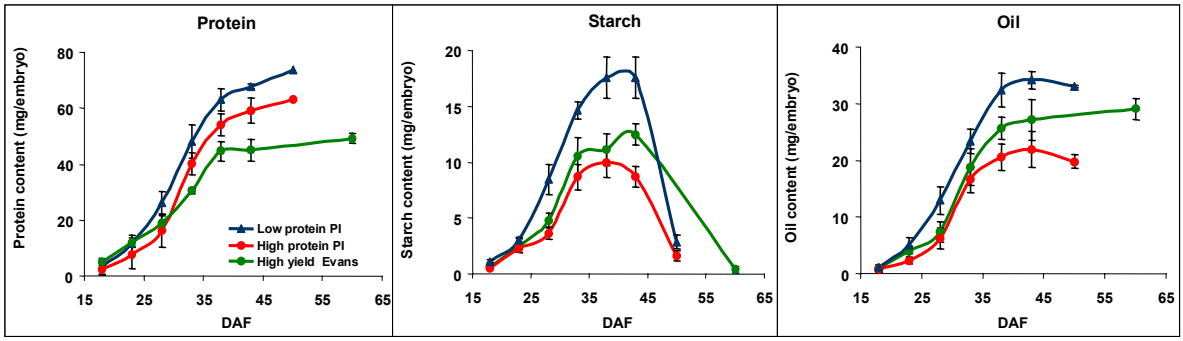


Figure 2. Protein, oil and starch content per embryo in high protein PI, low protein PI, and Evans during seed development.
Values are means of three or more replications. Bars represent standard error. Protein content is always increasing over seed development; oil content increases from the beginning and reaches its peak at 40-45 DAF, then decreases very fast, disappears from seed when seed is mature; oil accumulation is similar to that of protein's, but it reaches its peak at about 45DAF and maintains its high accumulation in mature seed. (From Dr. Westgate Lab)

Northern probe *ACLA* sequence (EST Genbank Accession number AI461199, 541 bp)

TTGTAACAGA CACCTGCAAA CCATCTCCAA CAATCGAGGA ATGCTAAAGA TACCTCTAAA
CATTCGCGCG CAACGCCATC CTCCATGGAG GAGGAAAGCC ACCGAGTCTT TCATTCATGG
ATTAGATGAA AAGACAAGTG CATCCTTAAA ATTCACAGTG TTGAACCCAA AGGGCCGAAT
TTGGACAATG GTAGCTGGGG GAGGTGCTAG TGTCATTTAT GCTGATACGG TAGGAGATCT
TGGCTTTGCA TCTGAGCTTG GAAACTATGC TGAATACAGT GGTGCACCCA AAGAAGAGGA
GGTCTTGCAG TACGCCAGAG TTGTAATTGA TTGTGCAACT GCAAACCCTG ATGACCAGAA
GAGAGCTCTT GTGATAGGAG GAGGCATAGC CAACTTTACT GATGTTGCTG CCACATTTAG
TGGTATAATT CGGGCATTGA AGGAGAAGGA GTCAAAATTG AAAGCAGCAA GGATGCACAT
NNCTTGGAGG AGAAGGGGTC CCAACTACCA GAAAGGTCTA GCTTTGATGC GAGCGCTTGG
A

Northern probe *ACLB* sequence (EST Genbank Accession number AI461059, 409 bp)

GCGCTTAGTC CTTATGAGTT TGTTGAAAGT ATGAAGAAGA AGGGAATTCG
TGTGCCAGGA ATAGGGCACA GGATCAAGAA TAGGGACAAC AAAGATAAGA
GAGTTGAGCT GCTACAGAAG TTTGCACGCA CACATTTTCC TTCTGTGAAA
TACATGGAAT ATGCTGTTCA AGTTGAGACC TACACGCTCA CAAAGGCAAA
TAATTTAGTT CTTAACGTAG ATGGTGCAAT TGGATCTCTT TTCTTGGATC
TTCTTGCTGG TAGTGGAATG TTCACCAAAC AAGAGATTGA TGAAATTGTG
GAGATTGGCT ATCTGAATGG CCTCTTTGTG CTGGCACGCT CCATTGGTCT
GATTGGGCAC ACCTTTGACC AAAAGCGATT GAAGCAACCA CTTTACCGTC
ACCCATGGG

Figure 3. Sequences of Northern probes *ACLA* and *ACLB*.

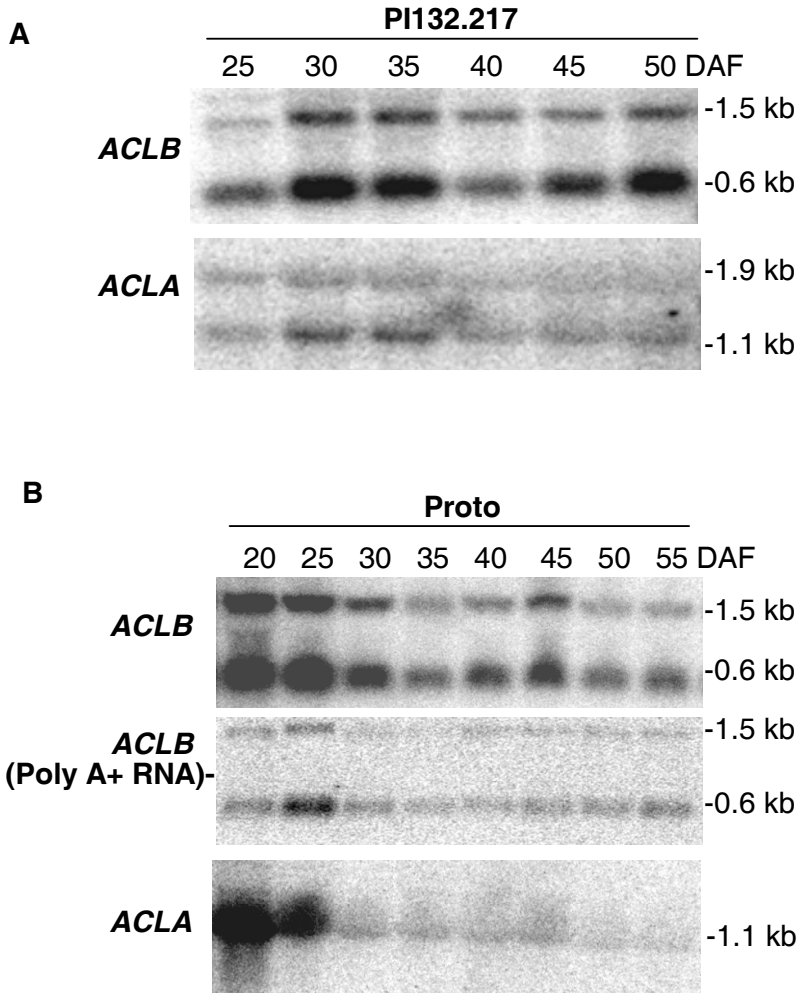


Figure 4. The RNA accumulation patterns of *ACLA* and *ACLB* are consistent with accumulation of oil.
Northern blots of PI132.217 (A) and proto (B) developing seeds RNAs. The RNA accumulation of *ACLB* from poly A+ RNA samples is similar to that from the total RNA, which proves the consistency of total RNA and poly A+ RNA so that we can compare the cDNA microarray data from poly A+ RNA with the Affymetrix microarray data. Probes for northern analysis are from EST Genbank Accession number AI461199 (*ACLA*) which is 541 bp; and from EST Genbank Accession number AI461059 (*ACLB*) which is 409 bp.

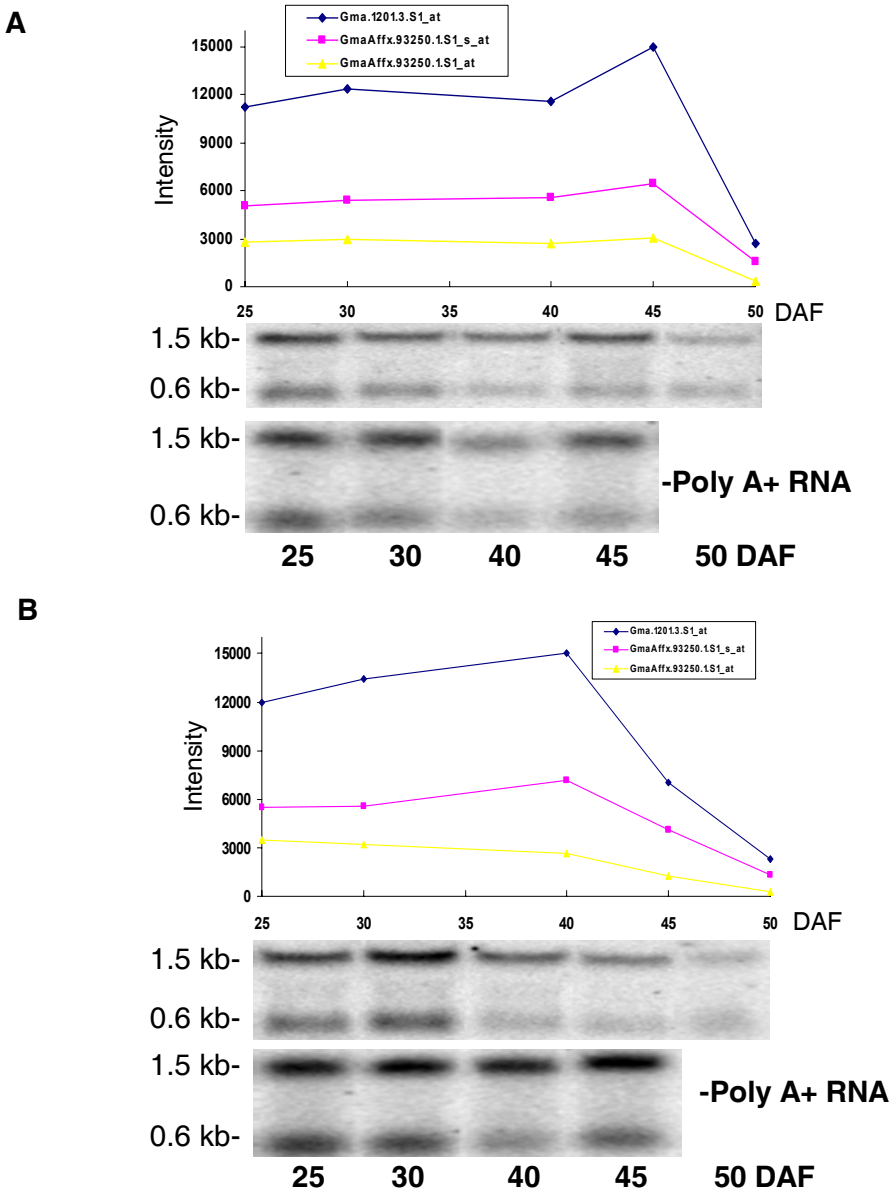


Figure 5. The RNA accumulation patterns for *ACLB* of Affymetrix microarray data for *Glycine Max* variety Evans in developing seeds is confirmed by northern blot analyses. Accumulation of *ACLB* RNA. Total RNAs were used unless poly A+ RNA is specified. (A) Replicate 1. (B) Replicate 2. Northern probe *ACLB* hybridizes with 2 targets: Gma.1201.3.S1_at (blue line, ~1.5 Kb) and GmaAffx.93250.1.S1_at (yellow line, ~0.6 Kb). Northern blots from poly A+ RNA also confirm that.

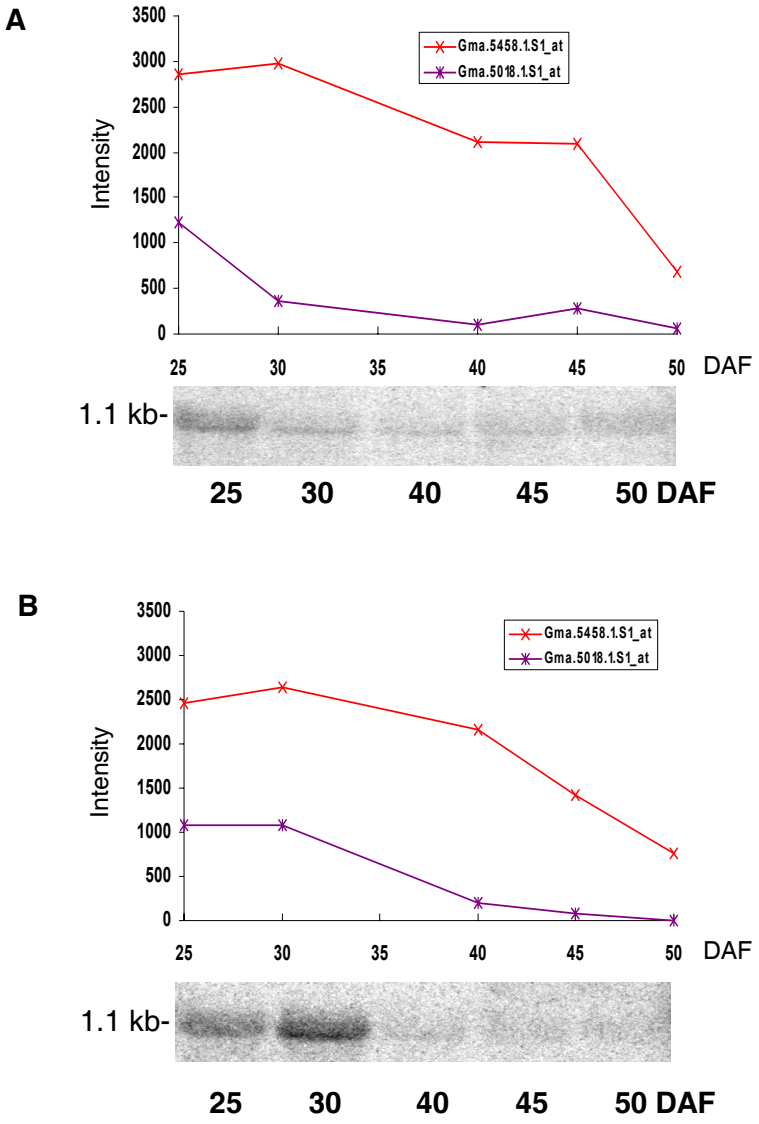


Figure 6. The RNA accumulation patterns for *ACLA* of Affymetrix microarray data for *Glycine Max* variety Evans in developing seeds is confirmed by northern blot analyses. Accumulation of *ACLA* RNA. Total RNAs were used. (A) Replicate 1. (B) Replicate 2. Northern probe *ACLA* hybridizes with Gma.5018.1.S1_at (purple line, ~1.1 Kb).

Table 1. Information about soybean seeds, RNA samples and Microarray experiments.

Experiment	Genotype		Stages	Replicates	RNA	Poly A^+ RNA	cDNA microarray	Affymetrix microarray
Evans 2000 Field	1	Evans	5	2	10	10	Y (10 $PolyA^+$)	Y (10 RNA)
PI, Proto, Evans 2000 GC	3	PI132.217	8	1	8	8	N	N
		Proto	10	1	10	10		
		Evans	8	1	8	N		
Temperature GC1 2001 (HTMTLT)	3	PI132.217	4 or 3	3	32	N	N	N
		Proto	4 or 3	3				
		Evans	4 or 3	3				
Evans, PI 2001 Field	1	Evans	6	3	18	N	N	N
		PI438.472	5	3	5			
		PI153.296	5	3	5			
Temperature GC2 2001 (HTMTLT)	3	PI132.217	4	3	39	N	N	N
		Proto	4	3				
		Evans	4	3		18	Y (18 Poly A^+, failed)	N
BC3 2002 Field	7	PI438.472	6	3	N	N	N	N
		PI153.296	5	3				
		Evans	7	3				
		BC3#2	7	1				
		BC3#14	7	2				
		BC3#33	7	2				
		BC3#45	7	2				
High Protein BC3 2003 GC	5	PI153.296	5	3	75	N	Y (72 RNA)	Y (75 RNA)
		Evans	5	3				
		BC3#33	5	3				
		BC3#93	5	3				
		BC3#128	5	3				

"GC" stands for growth chamber. "HTMTLT" stands for high temperature (35/27^{o}C), medium temperature (27/20^{o}C), and low temperature (20/12^{o}C).

Table 2**.** The northern blot probe information.

EST		Match to Soy Chip Probe		Soy Chip Probe ID	Size of Predicted SoyGene Transcript (bp)	Match to Arabidopsis Gene		Closest Arabidopsis Gene
Name	Size (bp)	Expect	Identities			Expect	Identities	
ACLA	541	0	429/444 (96%)	Gma.5458.1.S1_at	1900	0	617/757 (81%)	*AtACLA1*
ACLA	541	0	352/409 (86%)	Gma.5018.1.S1_at	1023	1E-68	486/606 (80%)	*AtACLA1*
ACLA	409	0	399/406 (98%)	Gma.1201.3.S1_at	1433	0	834/1033 (80%)	*AtACLB2*
ACLA	409	0	338/345 (97%)	GmaAffx.93250.1.S1_s_at	550	4E-49	428/538 (79%)	*AtACLB2*
ACLA	409	0	338/345 (97%)	GmaAffx.93250.1.S1_at	550	4E-49	428/538 (79%)	*AtACLB2*

The northern blot probe *ACLA* (used for Figure 2) matches 2 Soybean Affymetrix chip probe IDs with size ~1.9 kb and ~1 kb; probe *ACLB* (used for Figure 2) matches 3 soybean Affymetrix chip probe IDs with size ~1.5 kb and ~0.6 kb. Their closest matched genes in Arabidopsis are *AtACLB2* and *AtACLA1*. Sequence comparisons were done in BarleyBase (http://cereals.vrac.iastate.edu/contig_soybean.php).

APPENDIX C. SUPPLEMENTAL DATA FOR CHAPTER 1

Table S1. Genes of interest in this dissertation.

Gene	Locus ID	Biochemical function	Subcellular localization	Reference
GBSS	At1g32900	starch synthase	plastid	Ball et al., 1998, Trends Plant Sci 3: 462; Tenorio et al., 2003, Plant Mol Biol 51: 949
SS1	At5g24300	starch synthase, *AtSS1*	plastid	Delvallé et al., 2005, Plant J 43: 398
SS2	At3g01180	starch synthase, *AtSS2*	plastid	Smith et al., 1997, Annu Rev Plant Physiol Plant Mol Biol 48: 67
SS3	At1g11720	starch synthase, *AtSS3*	plastid	Zhang et al., 2005, Plant Physiol 138: 663
SS4	At4g18240	starch synthase, *AtSS4*	plastid (ChloroP)	Smith et al., 1997, Annu Rev Plant Physiol Plant Mol Biol 48: 67
SS6	At5g65685	starch synthase, *AtSS6*	plastid	Smith et al., 1997, Annu Rev Plant Physiol Plant Mol Biol 48: 67
BE1	At3g20440	starch branching enzyme	plastid	Dumez et al., 2006, Plant Cell 18: 2694
BE2	At5g03650	starch branching enzyme, *sbe2.2*	plastid	Dumez et al., 2006, Plant Cell 18: 2694
BE3	At2g36390	starch branching enzyme, *sbe2.1*	plastid	Dumez et al., 2006, Plant Cell 18: 2694
ISA1	At2g39930	isoamylase-type debranching enzyme, *ISO1*	plastid	Hussain et al., 2003, Plant Cell 15: 133; Delatte et al., 2005, Plant J 41: 815-830
ISA2	At1g03310	isoamylase-type debranching enzyme, *ISO2*	plastid	Hussain et al., 2003, Plant Cell 15: 133; Delatte et al., 2005, Plant J 41: 815-830
ISA3	At4g09020	isoamylase-type debranching enzyme, *ISO3*	plastid	Delatte et al., 2006, J Biol Chem 281: 12050
PU1	At5g04360	pullulanase-type starch debranching enzyme, limit dextrinase	plastid	Wattebled et al., 2005, Plant Physiol 138: 184
AAM1	At1g76130	alpha amylase, *AtAAM1*	non-plastidial	Yu et al., 2005, J Biol Chem 280: 9773
AAM2	At4g25000	alpha amylase, *AtAAM2*	non-plastidial	Yu et al., 2005, J Biol Chem 280: 9773
AAM3	At1g69830	alpha amylase, *AtAAM3*	plastid	Yu et al., 2005, J Biol Chem 280: 9773
DPE1	At5g64680	disproportionating enzyme, *DE1*	plastid	Critchley et al., 2001, Plant J 26: 89
DPE2	At2g40840	disproportionating enzyme, *DE2*, D-enzyme	cytosol	Lu et al., 2004, Planta 218: 466
PHS1	At3g29320	phosphorylase, starch phosophorylase, *PHO1*	plastid	Zeeman et al., 2004, Plant Physiol 135: 849
PHS2	At3g46970	phosphorylase, starch phosophorylase, *PHO2*	cytosol	Lu et al., 2006, Plant Physiol 142: 878
BAM1	At4g15210	beta amylase, *AtBAM1*, *ATb-Amy*, *RAM1*	?	Smith et al., 2005, Annu Rev Plant Biol 56: 73
BAM2	At2g32290	beta amylase, *AtBAM2*	?	Smith et al., 2005, Annu Rev Plant Biol 56: 73
BAM3	At3g23920	beta amylase, *AtBAM3*	plastid	Sparla et al., 2006, Plant Physiol 141: 840; Kaplan and Guy, 2004, Plant Physiol 135: 1674
BAM4	At2g45880	beta amylase, *AtBAM4*	?	Smith et al., 2005, Annu Rev Plant Biol 56: 73
BAM5	At5g45300	beta amylase, *AtBAM5*	?	Smith et al., 2005, Annu Rev Plant Biol 56: 73
BAM6	At5g55700	beta amylase, *AtBAM6*	?	Smith et al., 2005, Annu Rev Plant Biol 56: 73
BAM7	At4g00490	beta amylase, *AtBAM7*	plastid (ChloroP)	Smith et al., 2005, Annu Rev Plant Biol 56: 73
BAM8	At4g17090	beta amylase, *ct-BMY*, *AtBAM8*	plastid	Jung et al., 2003, Plant Mol Biol 52: 553; Lao et al., 1999, Plant J 20: 519; Kaplan and Guy, 2005, Plant J 44: 730
BAM9	At5g18670	beta amylase, *AtBAM9*	?	Smith et al., 2005, Annu Rev Plant Biol 56: 73

GWD1	At1g10760	Alpha-glucan water dikinase, *GWD*, *R1*, *SEX1*	plastid	Yu et al., 2001, Plant Cell 13: 1907; Yano et al., 2005, Plant Physiol 138: 837
GWD3	At5g26570	glucan water dikinase	plastid	Baunsgaard et al., 2005, Plant J 41: 595
SEX4	At3g52180	phosphatase/kinase	plastid	Fordham-Skelton et al., 2002, Plant J 29: 705; Niittylä et al., 2006, J Biol Chem 281: 11815
MEX1	At5g17520	Maltose exporter, Maltose transporter, *MT*	plastid	Niittylä et al., 2004, Science 303: 87
TPT1	At5g46110	Triose phosphate/Pi translocator, *TPT*	plastid	Walters et al., 2004, Plant Physiol 135: 891
GPT1	At5g54800	Glucose 6-phosphate/Pi transporter 1, *G6PT1*	plastid	Niewiadomski et al., 2005, Plant Cell 17: 760
GPT2	At1g61800	Glucose 6-phosphate/Pi transporter 2	plastid	Niewiadomski et al., 2005, Plant Cell 17: 760
GLT1	At5g16150	Glucose transporter, *GT*	plastid	Trethewey and ap Rees, 1994, Biochem J 301: 449
APL1	At5g19220	ADP-glucose pyrophosphorylase large subunit 1, *AGP*, *ADG2*	plastid	Crevillén et al., 2005, J Biol Chem 280: 8143
APL2	At1g27680	ADP-glucose pyrophosphorylase large subunit 2, *AGP*	?	Crevillén et al., 2005, J Biol Chem 280: 8143
APL3	At4g39210	ADP-glucose pyrophosphorylase large subunit 3, *AGP*	?	Crevillén et al., 2005, J Biol Chem 280: 8143
APL4	At2g21590	ADP-glucose pyrophosphorylase large subunit 4, *AGP*	?	Crevillén et al., 2005, J Biol Chem 280: 8143
APS1	At5g48300	ADP-glucose pyrophosphorylase small subunit 1, *AGP*, *ADG1*	plastid	Hendriks et al., 2003, Plant Physiol 133: 838
APS2	At1g05610	ADP-glucose pyrophosphorylase small subunit 2, *AGP*	plastid	Hendriks et al., 2003, Plant Physiol 133: 838
HXK1	At4g29130	Hexokinase 1	?	Moore et al., 2003, Science 300: 332
HXK2	At2g19860	Hexokinase 2	?	Frommer et al., 2003, Science 300: 261
HXK3	At1g50460	Hexokinase 3	?	Frommer et al., 2003, Science 300: 261
HXK4	At1g47840	Hexokinase 4	plastid (ChloroP)	Frommer et al., 2003, Science 300: 261
HXK5	At3g20040	Hexokinase 5	plastid (ChloroP)	Frommer et al., 2003, Science 300: 261
INV1	At2g36190	Invertase	?	Tang et al., 1996, Planta 198: 17
UGP	At3g03250	UDP-glucose pyrophosphorylase	?	Kleczkowski, 1996, Trends Plant Sci 1: 363
SPP1	At1g51420	Sucrose phosphate phosphatase 1	?	Lunn and REEs, 1990, Biochem J 267: 739
SPP2	At3g52340	Sucrose phosphate phosphatase 2	?	Lunn and REEs, 1990, Biochem J 267: 739
SPP4	At2g25780	Sucrose phosphate phosphatase 4	?	Lunn and REEs, 1990, Biochem J 267: 739
SPP3	At2g35800	Sucrose phosphate phosphatase 3	?	Lunn and REEs, 1990, Biochem J 267: 739
PGM1	At5g51820	Phosphoglucomutase	plastid	Caspar et al., 1985, Plant Physiol 79: 11

APPENDIX D. SUPPLEMENTAL DATA FOR CHAPTER 2

Table S1. Isoamylase-type debranching enzyme genes are differently expressed in Arabidopsis plants.

Tissue		***ISA1***	***ISA2***	***ISA3***	***PU1***
Root	Tip	-	+	+	N/D
	Vasculature	-	+(Old)	Less(old)	N/D
	Epidermis/Cortex	-	++	Less	N/D
	Root surface	+(Young)	-	-	N/D
	Root columella	-	+	+	N/D
Cotyledon	Petiole	Part	+	-	N/D
	Vasculature	Part	+	Less	N/D
	Mesophyll	+++	++	+	N/D
	Hydathode	-	+	+	N/D
Leaf	Shoot meristem	+	+	-	N/D
	Petiole	Part	+	-	N/D
	Vasculature	Part	+	Less(old)	N/D
	Mesophyll	+++	++	Less	N/D
	Hydathode	-	+	+	N/D
	Trichome	-	Base	Base(less)	N/D
Cauline leaf	Petiole	Part	+	-	N/D
	Vasculature	Part	+	Less	N/D
	Mesophyll	++	++	Less	N/D
	Hydathode	-	+	+	N/D
Stem	Vasculature	+	+	Less	N/D
	Cortex	+	+	+	N/D
	Epidermal	-	-	-	N/D
Flower	Pedicle	+	+	+	N/D
	Receptacle	-	+	-	N/D
	Sepal	+	++	++	N/D
	Petal	-	-	-	N/D
	Filament	+	+	-	N/D
	Anther	-	-	-	N/D
	Pollen	-	+	-	N/D
	Ovary wall	+	Less	+	N/D
	Style	-	+	+	N/D
	Stigma	+	Less	Less	N/D
Silique	Wall	+	Less	Less	N/D
	Seed/Embryo	Less	+	Less	N/D

DBE promoter::GUS expression under continuous light. More "+" means stronger signal. *PU1* expression is not detected (N/D) under continuous light.

Table S2. Genes that are highly correlated to all DBEs in RNA accumulation in MetaOmGraph.

Gene	*ISA1*	*ISA2*	*ISA3*	*PU1*	Locus ID
ISA2	0.74	1.00	0.58	0.73	AT1G03310
PU1	0.59	0.73	0.53	1.00	AT5G04360
DPE1	0.72	0.64	0.56	0.59	AT5G64860
MEX1	0.68	0.75	0.52	0.79	AT5G17520
PHS1	0.66	0.78	0.53	0.66	AT3G29320
DPE2	0.62	0.80	0.79	0.65	AT2G40840
SEX4	0.61	0.72	0.82	0.66	AT3G52180
PHS2	0.58	0.73	0.81	0.71	AT3G46970
AAM3	0.54	0.63	0.53	0.72	AT1G69830
GWD1	0.52	0.67	0.80	0.70	AT1G10760
GWD3	0.52	0.62	0.76	0.72	AT5G26570
sugar transporter	0.52	0.70	0.51	0.60	AT4G04750
BAM7	0.50	0.60	0.51	0.68	AT4G00490
ATP synthase protein I	0.68	0.75	0.51	0.83	AT2G31040
ABC1 family protein	0.64	0.69	0.50	0.78	AT5G05200
nocturnin-related, mitochondrion	0.63	0.75	0.52	0.81	AT3G18500
expressed protein, chloroplast	0.61	0.66	0.51	0.66	AT1G67660
thiamine biosynthesis	0.60	0.73	0.61	0.80	AT2G29630
AMP-binding protein	0.58	0.70	0.54	0.80	AT4G14070
expressed protein, chloroplast	0.58	0.67	0.54	0.74	AT3G15840
small heat shock family protein, mitochondrion	0.57	0.67	0.61	0.70	AT1G06460
acid phosphatase class B family protein	0.57	0.63	0.55	0.58	AT2G39920
targeted to thylakoid membrane.; stress-responsive protein	0.56	0.62	0.53	0.72	AT1G29390
expressed protein, unknown	0.54	0.66	0.60	0.59	AT1G68600
carbohydrate metabolism, expressed protein	0.52	0.58	0.50	0.61	AT4G32340
translational elongation, expressed protein,	0.51	0.60	0.50	0.61	AT1G13930
transporter activity, chloroplast, apolipoprotein D-related	0.51	0.63	0.53	0.79	AT3G47860
nucellin protein	0.51	0.63	0.65	0.61	AT4G33490

Table S3. Genes that are highly correlated to *ISA3* and GWD genes in RNA accumulation in MetaOmGraph.

Gene	ISA3	GWD1	GWD3	Locus ID
ISA3	1.00	0.80	0.76	AT4G09020
GWD1	0.80	1.00	0.85	AT1G10760
GWD3	0.76	0.85	1.00	AT5G26570
PHS2	0.81	0.92	0.83	AT3G46970
PHS1	0.53	0.67	0.64	AT3G29320
DPE2	0.79	0.85	0.74	AT2G40840
DPE1	0.56	0.67	0.68	AT5G64860
SEX4	0.82	0.82	0.81	AT3G52180
BE3	0.74	0.68	0.63	AT2G36390
ISA2	0.58	0.67	0.62	AT1G03310
PU1	0.53	0.70	0.72	AT5G04360
BAM2	0.56	0.69	0.72	AT2G32290
AAM3	0.53	0.79	0.72	AT1G69830
MEX1	0.52	0.69	0.74	AT5G17520
ACD32.1 small Hsp	0.61	0.82	0.74	AT1G06460
Hsp	0.64	0.74	0.74	AT5G23240
LTI6A	0.63	0.74	0.61	AT4G30650
COR15B	0.62	0.72	0.58	AT2G42530
COR15A	0.58	0.66	0.52	AT2G42540
COR314-TM2	0.53	0.70	0.64	AT1G29390
COR414-TM1	0.61	0.71	0.59	AT1G29395
KIN1	0.64	0.67	0.56	AT5G15960
ATGRP7 circadian	0.56	0.58	0.60	AT2G21660
APRR3	0.60	0.58	0.68	AT5G60100
BEL1	0.50	0.51	0.53	AT5G41410
S-receptor protein kinase	0.54	0.67	0.64	AT1G65800
ATP synthase	0.51	0.76	0.75	AT2G31040
Phosphatase	0.55	0.68	0.62	AT2G39920
acid phosphatase class B family protein	0.52	0.75	0.68	AT3G18500
Constans	0.53	0.68	0.69	AT1G07050
member of Arabidopsis OEP16 family	0.62	0.77	0.71	AT2G28900
thiC family protein	0.61	0.75	0.79	AT2G29630
apolipoprotein D-related	0.53	0.68	0.73	AT3G47860
Plastidic acyl activating enzyme	0.54	0.68	0.73	AT4G14070
similar to nucellin protein	0.65	0.74	0.69	AT4G33490
ABC1 family protein	0.50	0.75	0.72	AT5G05200
disease resistance protein	0.50	0.67	0.64	AT5G40060
zinc finger (B-box type) family protein	0.55	0.65	0.64	AT5G48250
expressed protein	0.50	0.68	0.52	AT1G13930
expressed protein	0.51	0.62	0.71	AT1G17665
expressed protein	0.51	0.69	0.64	AT1G67660
expressed protein	0.60	0.68	0.65	AT1G68600
expressed protein	0.52	0.62	0.54	AT2G34090
expressed protein	0.55	0.64	0.78	AT2G47390
expressed protein	0.54	0.75	0.69	AT3G15840
expressed protein	0.50	0.64	0.52	AT3G50960
expressed protein	0.58	0.64	0.59	AT4G16146
expressed protein	0.50	0.71	0.70	AT4G32340

Table S4. Cold and light responsive promoter motifs exist commonly among the starch metabolism genes which are not only highly correlated to slow cold responsive genes but also have obvious changes in response to cold.

Motif/Gene	*PHS2*	*BAM8*	*BE3*	*T6PP*	*BAM7*	*ISA3*	*SEX4*	*GWD1*	*SS1*	*GWD3*	*ADG1*	*ADG2*	*AAM3*	*DPE2*	*PGM1*	P-value
DREB1A/CBF3																0.0115
DRE core motif																0.0601
MYB1AT																0.7203
MYB4 binding site																0.4853
ABRE-like binding site																0.0017
ACGTABRE A2OSEM																<0.001
ABF																<0.001
ABRE																<0.001
LEAFYATAG																0.3194
LTRE promoter motif																0.0796
AtMYC2 BS in RD22																0.0275
ABREATRD22																0.015
CArG promoter motif																0.1791
Evening Element																<0.0001
CACGTGMOTIF																<0.001
GBOXLERBCS																<0.0001
ARF binding site																0.0377
CCA1 binding site																0.3281
CARGCW8GAT																0.8409
GBF1/2/3 BS in ADH1																0.0038
Ibox promoter motif																0.0085
BoxII promoter motif																0.8286

Promoter motif analysis (from Athena) of selected genes from the 25 starch metabolic genes that are co-expressed with cold responsive genes. Cold responsive promoter motifs are labeled light blue; light responsive promoter motifs are labeled lilac. Genes with increased transcript accumulation in cold are labeled light blue.

APPENDIX E. SUPPLEMENTAL DATA FOR CHAPTER 3

Table S1. Genes with most significantly changed transcripts over the SD diurnal cycle in WT (FDR is controlled at the level of 0.00031). Supplement to Figure 3. Carbohydrate metabolic genes, genes involved in transcription, genes responded to cold.

Cluster 1

Locus ID	p-value	q-value	Annotation	MapMan Category
At4g22570	3.30E-06	0.0002876	adenine phosphoribosyltransferase	nucleotide metabolism.salvage
At1g08460	5.46E-08	3.36E-05	histone deacetylase family protein	RNA.regulation of transcription.HDA
At1g05890	7.59E-07	0.0001203	zinc finger protein	RNA.regulation of transcription.unclassified
At1g49200	9.76E-07	0.0001345	zinc finger (C3HC4-type RING finger) family protein	RNA.regulation of transcription.unclassified
At3g17590	3.54E-06	0.000305	transcription regulatory protein SNF5	RNA.regulation of transcription.Chromatin Remodeling Factors
At5g35330	3.23E-06	0.000287	methyl-CpG-binding domain-containing protein	RNA.regulation of transcription.Methyl binding domain proteins
At5g22000	5.80E-07	0.0001053	zinc finger (C3HC4-type RING finger)	RNA.regulation of transcription.unclassified
At1g47270	3.12E-07	8.06E-05	F-box family protein / tubby family protein	protein.degradation.ubiquitin.E3.SCF.FBOX; cell.organisation
At4g27130	7.06E-07	0.0001164	Translation initiation factor SUI1	protein.synthesis.initiation
At1g64040	9.69E-07	0.0001345	phosphoprotein Ser/Thr phosphatase	protein.postranslational modification
At2g25070	2.12E-06	0.0002243	protein phosphatase 2C	protein.postranslational modification
At3g02740	8.29E-07	0.0001225	aspartyl protease	protein.degradation.aspartate protease
At5g17290	1.01E-07	4.17E-05	Autophagy protein ATG5	protein.degradation.autophagy
At1g69295	5.51E-07	0.0001027	beta-1,3-glucanase-related	misc.beta 1,3 glucan hydrolases
At1g47960	2.24E-06	0.0002326	invertase/pectin methylesterase inhibitor	cell wall.pectin*esterases.PME; misc.invertase/pectin methylesterase inhibitor family protein
At4g14900	4.66E-07	9.72E-05	hydroxyproline-rich glycoprotein	cell wall.cell wall proteins.HRGP
At4g15780	1.10E-08	1.39E-05	synaptobrevin-related	cell. vesicle transport
At4g32150	1.35E-06	0.0001678	synaptobrevin family protein	cell. vesicle transport
At2g30540	3.06E-06	0.0002776	GLUTAREDOXIN	redox.glutaredoxins
At3g01520	1.30E-06	0.0001661	universal stress protein	stress.abiotic.unspecified
At1g75750	3.29E-06	0.0002876	gibberellin-regulated protein 1 (GASA1)	hormone metabolism.gibberelin.induced-regulated-responsive-activated
At1g05810	1.20E-06	0.000156	Ras-related protein (ARA-1)	signalling.G-proteins
At1g16920	2.21E-06	0.0002316	Ras-related GTP-binding protein	signalling.G-proteins
At1g74880	6.42E-07	0.000109	Ndh complex	not assigned.unknown
At2g39000	6.21E-07	0.000109	N-acetyltransferase	not assigned.no ontology
At3g54120	9.74E-07	0.0001345	reticulon family protein	not assigned.no ontology
At5g56750	3.42E-07	8.22E-05	Ndr family protein	not assigned.no ontology
At1g30880	1.62E-06	0.0001919	expressed protein	not assigned.unknown
At1g68440	8.11E-07	0.0001225	expressed protein	not assigned.unknown
At3g05760	1.79E-06	0.0002046	expressed protein	not assigned.unknown
At4g18740	1.17E-06	0.0001529	expressed protein	not assigned.unknown
At3g55420	2.12E-06	0.0002243	expressed protein	not assigned.unknown
At5g11950	2.60E-06	0.000256	expressed protein	not assigned.unknown
At5g42680	9.09E-10	6.26E-06	expressed protein	not assigned.unknown
At1g58032	3.08E-06	0.0002776	unknown	

Cluster 2

Locus ID	p-value	q-value	Annotation	MapMan Category
At1g70000	4.93E-07	9.78E-05	DNA-binding family protein	RNA.regulation of transcription.MYB-related transcription factor family

At2g20570	7.67E-07	0.0001203	golden2-like transcription factor	RNA.regulation of transcription.G2-like transcription factor family, GARP
At3g09600	2.12E-09	6.89E-06	myb family transcription factor	RNA.regulation of transcription.MYB-related transcription factor family
At3g02830	6.79E-08	3.59E-05	zinc finger (CCCH-type) family protein	RNA.regulation of transcription.C2H2 zinc finger family
At4g00050	4.84E-08	3.18E-05	basic helix-loop-helix (bHLH) family protein	RNA.regulation of transcription.bHLH,Basic Helix-Loop-Helix family
At5g02840	3.36E-09	7.64E-06	myb family transcription factor	RNA.regulation of transcription.MYB-related transcription factor family
At5g44190	9.94E-07	0.0001362	myb family transcription factor (GLK2)	RNA.regulation of transcription.G2-like transcription factor family, GARP
At5g15830	1.05E-06	0.000143	bZIP transcription factor family protein	RNA.regulation of transcription.bZIP transcription factor family
At5g15850	2.99E-06	0.0002752	zinc finger protein CONSTANS-LIKE 1	RNA.regulation of transcription.C2C2(Zn) CO-like, Constans-like zinc finger family; development.unspecified
At2g15580	1.31E-06	0.0001665	zinc finger (C3HC4-type RING finger) family protein	RNA.regulation of transcription.C3H zinc finger family
At1g60000	1.25E-06	0.0001615	ribonucleoprotein	RNA.regulation of transcription.unclassified; RNA.RNA binding
At4g27000	7.66E-07	0.0001203	RNA-binding protein 45	RNA.regulation of transcription.unclassified
At5g28020	4.11E-08	2.83E-05	cysteine synthase	amino acid metabolism.synthesis.serine-glycine-cysteine group.cysteine.OASTL
At2g16500	6.22E-07	0.000109	arginine decarboxylase	amino acid metabolism.degradation.glutamate family.arginine; polyamine metabolism
At5g27360	9.55E-07	0.0001345	sugar-porter family protein 2	transporter.sugars
At2g47490	1.65E-09	6.26E-06	mitochondrial substrate carrier family protein	transport.metabolite transporters at the mitochondrial membrane
At3g21390	4.23E-07	9.16E-05	mitochondrial substrate carrier family protein	transport.metabolite transporters at the mitochondrial membrane
At5g47560	3.54E-07	8.22E-05	malate/fumarate transporter.	transport.unspecified cations
At2g26690	9.20E-10	6.26E-06	nitrate transporter	transport.nitrate
At1g33110	5.40E-07	0.0001024	MATE efflux family protein	transport misc
At5g49730	1.51E-07	4.98E-05	ferric chelate reductase	metal handling.acquisition
At4g25700	1.07E-06	0.000144	beta-carotene hydroxylase	secondary metabolism.isoprenoids.carotenoids
At1g64500	7.63E-08	3.84E-05	glutaredoxin family protein	redox.glutaredoxins
At1g52540	2.21E-07	6.36E-05	protein kinase	signalling.receptor kinases; development.unspecified
At1g62960	1.65E-07	5.29E-05	1-aminocyclopropane-1-carboxylate synthase	hormone metabolism.ethylene.synthesis-degradation
At3g14770	3.60E-06	0.0003055	nodulin MtN3 family protein	development.unspecified
At1g07010	2.13E-06	0.0002243	calcineurin-like phosphoesterase family protein	misc.calcineurin-like phosphoesterase family protein
At1g13080	3.56E-08	2.61E-05	cytochrome P450 family protein	misc.cytochrome P450
At4g37550	1.35E-06	0.0001678	formamidase	misc.misc2
At3g25585	1.05E-07	4.17E-05	aminoalcoholphosphotransferase	misc.misc2
At2g41250	1.18E-07	4.19E-05	Haloacid dehalogenase-like hydrolase	not assigned.no ontology
At2g19650	5.17E-07	9.97E-05	DC1 domain-containing protein	not assigned.no ontology
At4g25830	1.09E-07	4.17E-05	integral membrane family protein	not assigned.no ontology
At5g62130	9.50E-07	0.0001345	Per1-like protein-related	not assigned.no ontology
At4g26850	1.42E-08	1.70E-05	ascorbate biosynthesis	not assigned.unknown
At5g19290	3.59E-06	0.0003055	esterase/lipase/thioesterase family protein	not assigned.no ontology
At4g26860	3.93E-09	8.13E-06	alanine racemase family protein	not assigned.no ontology
At3g11620	1.34E-06	0.0001678	expressed protein	not assigned.unknown
At3g60910	3.51E-07	8.22E-05	expressed protein	not assigned.no ontology
At1g62250	2.88E-06	0.0002718	expressed protein	not assigned.unknown
At2g04039	2.53E-06	0.0002523	expressed protein	not assigned.unknown
At1g29700	2.54E-06	0.0002523	expressed protein	not assigned.unknown

At3g15310	5.03E-09	8.80E-06	expressed protein	not assigned.unknown
At1g75180	1.38E-07	4.62E-05	expressed protein	not assigned.unknown
At2g33250	2.51E-06	0.0002515	expressed protein	not assigned.unknown
At4g24700	1.76E-07	5.48E-05	expressed protein	not assigned.unknown
At5g18040	1.28E-06	0.0001645	expressed protein	not assigned.unknown

Cluster 3

Locus ID	p-value	q-value	Annotation	MapMan Category
At2g21320	3.23E-07	8.16E-05	zinc finger (B-box type) family protein	RNA.regulation of transcription.C2C2(Zn) CO-like, Constans-like zinc finger family
At2g31380	2.40E-07	6.71E-05	B-box zinc finger protein	RNA.regulation of transcription.C2C2(Zn) CO-like, Constans-like zinc finger family
At1g01060	4.45E-07	9.48E-05	myb family transcription factor	RNA.regulation of transcription.MYB-related transcription factor family
At3g02380	7.55E-07	0.0001203	zinc-finger protein	RNA.regulation of transcription.C2C2(Zn) CO-like, Constans-like zinc finger family; development.unspecified
At3g21890	1.51E-06	0.0001837	zinc finger (B-box type) family protein	RNA.regulation of transcription.C2C2(Zn) CO-like, Constans-like zinc finger family
At4g38960	2.22E-06	0.0002316	zinc finger (B-box type) family protein	RNA.regulation of transcription.C2C2(Zn) CO-like, Constans-like zinc finger family
At5g59780	8.14E-07	0.0001225	myb family transcription factor (MYB59)	RNA.regulation of transcription.MYB domain transcription factor family
At2g26800	6.27E-07	0.000109	hydroxymethylglutaryl-CoA lyase	amino acid metabolism.degradation. branched-chain group.leucine; secondary metabolism.isoprenoids.mevalonate pathway
At1g55920	3.41E-06	0.0002949	Encodes a chloroplast/cytosol localized serine O-acetyltransferase	amino acid metabolism.synthesis.serine-glycine-cysteine group.cysteine.SAT; metal handling.binding, chelation and storage
At3g08730	1.46E-06	0.0001795	protein-serine kinase	protein.synthesis.misc ribososomal protein
At5g15950	3.26E-08	2.61E-05	adenosylmethionine decarboxylase	polyamine metabolism
At5g44530	8.33E-07	0.0001225	subtilase family protein	protein.degradation.subtilases
At2g22990	4.86E-07	9.78E-05	sinapoylglucose:malate sinapoyltransferase	protein.degradation.serine protease
At2g27420	4.79E-07	9.78E-05	cysteine proteinase	protein.degradation.cysteine protease
At2g24540	2.64E-06	0.0002566	kelch repeat-containing F-box family protein	protein.degradation.ubiquitin.E3.SCF.FBOX
At5g47970	2.54E-07	6.96E-05	nitrogen regulation family protein	N-metabolism.misc
At1g07180	3.84E-08	2.73E-05	NAD(P)H dehydrogenase	mitochondrial electron transport / ATP synthesis.NADH-DH.type II.internal matrix
At1g80760	2.13E-07	6.21E-05	major intrinsic family protein	transport.Major Intrinsic Proteins.NIP
At3g01550	3.26E-09	7.64E-06	AtPPT2	transport.metabolite transporters at the envelope membrane
At1g19450	2.52E-08	2.39E-05	integral membrane protein	transporter.sugars
At3g11900	1.95E-06	0.0002134	amino acid transporter family protein	transport.amino acids
At5g64840	1.75E-06	0.0002021	ABC transporter family protein	transport.ABC transporters and multidrug resistance systems
At2g46800	1.06E-06	0.0001435	member of Zinc transporter (ZAT) family	transport.metal
At1g58290	7.23E-08	3.74E-05	glutamyl-tRNA reductase	tetrapyrrole synthesis
At1g44446	2.88E-07	7.53E-05	chlorophyll a oxygenase (CAO) / chlorophyll b synthase	tetrapyrrole synthesis
At5g64930	7.59E-07	0.0001203	CPR5 protein	stress.biotic
At1g65870	4.00E-07	8.92E-05	disease resistance-responsive family protein	stress.biotic
At5g67030	2.72E-07	7.28E-05	zeaxanthin epoxidase	hormone metabolism.abscisic acid.synthesis-degradation
At5g54130	1.08E-07	4.17E-05	calcium-binding EF hand family protein	signalling.calcium
At1g32900	3.36E-09	7.64E-06	GBSS	major CHO metabolism.synthesis. starch.starch synthase

At2g22240	6.48E-08	3.59E-05	inositol-3-phosphate synthase isozyme 2	minor CHO metabolism.myo-inositol.InsP Synthases
At1g78510	1.98E-06	0.0002134	solanesyl diphosphate synthase (SPS)	secondary metabolism. isoprenoids. terpenoids
At1g07450	9.60E-07	0.0001345	tropinone reductase	secondary metabolism.N misc.alkaloid-like; misc.nitrilases, *nitrile lyases, berberine bridge enzymes, reticuline oxidases, troponine reductases
At2g34460	2.37E-08	2.34E-05	flavin reductase-related	not assigned.no ontology
At4g39510	1.37E-06	0.0001694	cytochrome P450 family protein	misc.cytochrome P450
At3g17040	3.00E-08	2.61E-05	RNA tetratricopeptide repeat-containing protein	not assigned.no ontology.pentatricopeptide (PPR) repeat-containing protein
At2g41010	1.81E-06	0.0002048	VQ motif-containing protein	not assigned.no ontology
At1g17870	6.38E-07	0.000109	S2P-like putative metalloprotease	not assigned.unknown
At4g15550	2.11E-07	6.21E-05	UDP-glucose:indole-3-acetate beta-D-glucosyltransferase	not assigned.no ontology
At3g61080	5.89E-08	3.37E-05	fructosamine kinase family protein	not assigned.no ontology
At5g43850	8.35E-07	0.0001225	acireductone dioxygenase	not assigned.no ontology
At5g15910	2.25E-06	0.0002326	dehydrogenase-related	not assigned.no ontology
At5g20380	2.01E-06	0.0002157	transporter-related	not assigned.no ontology
At2g41120	1.52E-08	1.73E-05	expressed protein	not assigned.unknown
At2g36630	2.94E-06	0.0002741	expressed protein	not assigned.unknown
At1g22850	2.69E-06	0.0002593	expressed protein	not assigned.unknown
At1g55960	2.38E-06	0.0002417	expressed protein	not assigned.no ontology
At1g69160	8.22E-07	0.0001225	expressed protein	not assigned.unknown
At3g12320	3.00E-06	0.0002752	expressed protein	not assigned.unknown
At1g16730	4.45E-09	8.43E-06	expressed protein	not assigned.unknown
At4g17840	1.02E-07	4.17E-05	expressed protein	not assigned.unknown
At4g27030	1.08E-06	0.0001445	expressed protein	not assigned.unknown
At3g54500	1.90E-06	0.0002108	expressed protein	not assigned.unknown
At3g56290	1.68E-06	0.000196	expressed protein	not assigned.unknown
At5g43180	1.53E-06	0.0001841	expressed protein	not assigned.unknown
At5g54120	7.76E-08	3.84E-05	expressed protein	not assigned.unknown
At5g59400	6.97E-07	0.0001157	expressed protein	not assigned.unknown
At5g61670	8.10E-07	0.0001225	expressed protein	not assigned.unknown
At5g12470	3.33E-07	8.22E-05	expressed protein	not assigned.unknown

Cluster 4

Locus ID	p-value	q-value	Annotation	MapMan Category
At5g35970	1.89E-07	5.66E-05	DNA-binding protein	DNA.unspecified
At1g06040	1.12E-06	0.000149	B-box zinc finger	RNA.regulation of transcription.C2C2(Zn) CO-like, Constans-like zinc finger family
At4g18390	2.94E-06	0.0002741	TCP family transcription factor	RNA.regulation of transcription.TCP transcription factor family
At4g30960	3.07E-06	0.0002776	CBL-interacting protein kinase 6	protein.postranslational modification
At5g45800	1.97E-06	0.0002134	leucine-rich repeat transmembrane protein kinase	protein.postranslational modification
At5g02180	8.62E-08	4.00E-05	amino acid transporter family protein	transport.amino acids
At2g29650	6.72E-08	3.59E-05	inorganic phosphate transporte	transport.phosphate
At4g13800	8.07E-07	0.0001225	permease-related	transport misc
At2g05070	1.16E-07	4.19E-05	chlorophyll A-B binding protein	PS.lightreaction.photosystem II.LHC-II
At2g36870	3.09E-06	0.0002776	xyloglucan:xyloglucosyl transferase	cell wall.modification
At1g15950	9.27E-08	4.17E-05	cinnamoyl CoA reductase	secondary metabolism.phenylpropanoids.lignin biosynthesis.CCR1
At2g39980	1.84E-06	0.0002072	transferase	secondary metabolism.flavonoids.anthocyanins
At4g32770	2.62E-06	0.0002566	Tocopherol cyclase	secondary metabolism. isoprenoids. tocopherol biosynthesis

At3g57020	4.84E-07	9.78E-05	strictosidine synthase	secondary metabolism.N misc.alkaloid-like
At4g13010	8.06E-08	3.90E-05	oxidoreductase	misc.oxidases - copper, flavone etc.
At1g10370	1.75E-06	0.0002021	glutathione S-transferase	misc.glutathione S transferases
At3g48460	7.73E-07	0.0001204	GDSL-motif lipase/hydrolase family protein	misc.GDSL-motif lipase
At5g42250	9.12E-07	0.0001305	alcohol dehydrogenase	misc.alcohol dehydrogenases
At1g18360	2.28E-06	0.0002336	hydrolase	not assigned.no ontology
At1g73480	3.06E-08	2.61E-05	hydrolase, alpha/beta fold family protein	not assigned.no ontology
At2g40400	2.85E-07	7.53E-05	expressed protein	not assigned.unknown

Cluster 5

Locus ID	p-value	q-value	Annotation	MapMan Category
At3g57150	2.67E-06	0.0002584	dyskerin	DNA.synthesis/chromatin structure
At4g04840	1.51E-06	0.0001837	methionine sulfoxide reductase domain-containing protein / SeIR domain-containing protein	RNA.regulation of transcription.putative DNA-binding protein
At3g56330	1.16E-06	0.0001529	N2,N2-dimethylguanosine tRNA methyltransferase	RNA.regulation of transcription.unclassified
At4g39260	7.31E-07	0.0001191	glycine-rich RNA binding protein CCR1	RNA.regulation of transcription.GRP
At1g11790	3.00E-06	0.0002752	prephenate dehydratase family protein	amino acid metabolism.synthesis.aromatic aa.phenylalanine
At2g31610	3.06E-06	0.0002776	40S ribosomal protein S3 (RPS3A)	protein.synthesis.misc ribososomal protein
At4g02930	3.71E-07	8.52E-05	elongation factor Tu	protein.synthesis.elongation
At4g33490	3.97E-07	8.92E-05	nucellin protein	protein.degradation
At1g69830	9.04E-07	0.0001301	AAM3	major CHO metabolism.degradation.starch.starch cleavage
At1g22650	3.21E-06	0.0002863	beta-fructofuranosidase	major CHO metabolism.degradation.sucrose.misc
At5g26340	3.23E-07	8.16E-05	hexose transporter	transporter.sugars
At1g22570	5.31E-08	3.36E-05	proton-dependent oligopeptide transport	transport.peptides and oligopeptides
At5g62360	1.57E-06	0.000188	invertase/pectin methylesterase inhibitor family protein, INV	misc.invertase/pectin methylesterase inhibitor family protein
At5g03350	3.32E-06	0.0002882	legume lectin family protein	misc.myrosinases-lectin-jacalin
At2g42530	9.63E-09	1.29E-05	cold-responsive protein / cold-regulated protein (cor15b)	stress.abiotic.cold
At1g20450	2.97E-06	0.0002752	dehydrin (ERD10)	stress.abiotic.unspecified
At1g56300	1.09E-07	4.17E-05	DNAJ heat shock N-terminal domain-containing protein	stress.abiotic.heat
At5g47220	2.83E-06	0.0002693	ERF (ethylene response factor) subfamily B-3 of ERF/AP2 transcription factor family	hormone metabolism.ethylene.signal transduction
At3g07390	1.88E-06	0.0002096	auxin-responsive protein	hormone metabolism.auxin.induced-regulated-responsive-activated
At3g16050	3.28E-06	0.0002876	stress-responsive protein	hormone metabolism.ethylene.induced-regulated-responsive-activated
At3g44860	2.63E-06	0.0002566	S-adenosyl-L-methionine:carboxyl methyltransferase	hormone metabolism.salicylic acid.synthesis-degradation
At5g39790	1.81E-06	0.0002048	5'-AMP-activated protein kinase beta-1 subunit-related	signalling.in sugar and nutrient physiology
At5g45750	3.12E-08	2.61E-05	Ras-related GTP-binding protein	signalling.G-proteins
At2g45560	3.56E-08	2.61E-05	cytochrome P450 family protein	misc.cytochrome P450
At2g34260	5.69E-08	3.37E-05	transducin family protein	not assigned.no ontology
At1g76590	3.27E-06	0.0002876	zinc-binding family protein	not assigned.no ontology
At3g17170	9.62E-09	1.29E-05	ribosomal protein S6 family protein	not assigned.no ontology
At3g18760	7.51E-09	1.21E-05	ribosomal protein S6 family protein	not assigned.no ontology
At3g27280	6.28E-07	0.000109	prohibitin	not assigned.no ontology
At4g34120	5.77E-07	0.0001053	CBS domain-containing protein	not assigned.no ontology
At5g38890	2.74E-06	0.000263	exoribonuclease-related	not assigned.no ontology
At5g47240	4.96E-07	9.78E-05	MutT/nudix family protein	not assigned.no ontology

At5g19750	1.03E-07	4.17E-05	peroxisomal membrane 22 kDa family protein	not assigned.no ontology
At1g75860	1.66E-06	0.0001946	expressed protein	not assigned.unknown
At2g22870	2.62E-07	7.09E-05	expressed protein	not assigned.unknown
At2g38790	1.81E-07	5.49E-05	expressed protein	not assigned.unknown
At1g10522	4.46E-07	9.48E-05	expressed protein	not assigned.unknown
At1g28395	4.59E-07	9.67E-05	expressed protein	not assigned.unknown
At1g48460	4.10E-07	9.05E-05	expressed protein	not assigned.unknown
At1g20220	2.50E-06	0.0002515	expressed protein	RNA.regulation of transcription.putative DNA-binding protein
At1g27930	5.48E-07	0.0001027	expressed protein	not assigned.unknown
At3g15000	3.54E-07	8.22E-05	expressed protein	not assigned.no ontology
At4g24175	7.33E-07	0.0001191	expressed protein	not assigned.unknown
At5g47455	2.30E-06	0.0002346	expressed protein	not assigned.unknown

Cluster 6

Locus ID	p-value	q-value	Annotation	MapMan Category
At1g34380	7.95E-09	1.21E-05	5'-3' exonuclease family protein	DNA.synthesis/chromatin structure
At3g27360	3.27E-06	0.0002876	histone H3	DNA.synthesis/chromatin structure.histone
At5g10390	2.44E-06	0.0002467	histone H3	DNA.synthesis/chromatin structure.histone
At1g05805	1.23E-07	4.30E-05	basic helix-loop-helix (bHLH) family protein	RNA.regulation of transcription.bHLH,Basic Helix-Loop-Helix family
At1g49230	1.10E-07	4.17E-05	zinc finger (C3HC4-type RING finger) family protein	RNA.regulation of transcription.unclassified
At1g19310	3.38E-07	8.22E-05	zinc finger (C3HC4-type RING finger) family protein	RNA.regulation of transcription.C3H zinc finger family
At5g50450	2.09E-06	0.0002232	zinc finger (MYND type) family protein	RNA.regulation of transcription.unclassified
At3g53460	8.90E-07	0.0001289	29 kDa ribonucleoprotein	RNA.regulation of transcription.unclassified; RNA.RNA binding
At2g21660	4.71E-11	1.07E-06	glycine-rich RNA-binding protein (GRP7)	RNA.RNA binding
At3g22200	1.80E-07	5.49E-05	4-aminobutyrate aminotransferase	amino acid metabolism.synthesis.central amino acid metabolism.GABA
At2g28900	2.28E-06	0.0002336	mitochondrial import inner membrane translocase subunit	transport.protein transport mitochondrial membrane (TIM/TOM)
At4g13510	1.08E-07	4.17E-05	ammonium transporter	transport.ammonium
At4g26670	1.13E-07	4.19E-05	mitochondrial import inner membrane translocase subunit	transport.protein transport mitochondrial membrane (TIM/TOM)
At5g11150	1.66E-06	0.0001946	synaptobrevin / vesicle-associated membrane protein	cell. vesicle transport
At1g76790	8.60E-07	0.0001254	O-methyltransferase	secondary metabolism.simple phenols
At4g38580	1.52E-06	0.0001839	heavy-metal-associated domain-containing protein	metal handling.binding, chelation and storage
At5g06690	6.36E-07	0.000109	thioredoxin family protein	redox.thioredoxin
At3g26740	1.17E-06	0.0001529	transcripts are differentially regulated at the level of mRNA stability at different times of day controlled by the circadian clock	signalling.light
At2g02100	1.51E-09	6.26E-06	plant defensin-fusion protein	stress.biotic
At1g53350	2.59E-06	0.000256	disease resistance protein	stress.biotic
At2g43530	3.60E-06	0.0003055	defensin-like (DEFL) family protein	stress.biotic
At4g30650	1.18E-07	4.19E-05	low temperature and salt responsive protein	stress.abiotic.cold; stress.abiotic.drought/salt
At4g30660	4.84E-07	9.78E-05	low temperature and salt responsive protein	stress.abiotic.cold; stress.abiotic.drought/salt
At4g13850	1.97E-06	0.0002134	glycine-rich RNA-binding protein (GRP2)	stress.abiotic.cold; RNA.regulation of transcription.unclassified; RNA.RNA binding
At2g33830	1.27E-09	6.26E-06	dormancy/auxin associated family protein	hormone metabolism.auxin.induced-regulated-responsive-activated
At5g61380	8.42E-08	3.99E-05	ABI3-interacting protein 1 (AIP1)	hormone metabolism.abscisic acid.signal transduction; signalling.receptor kinases

At3g62550	5.92E-08	3.37E-05	universal stress protein	hormone metabolism.ethylene.induced-regulated-responsive-activated
At2g39900	4.89E-08	3.18E-05	LIM domain-containing protein	development.unspecified
At2g16700	8.23E-07	0.0001225	actin depolymerizing factor 5 (ADF5)	cell.organisation
At3g05220	2.42E-07	6.71E-05	heavy-metal-associated domain-containing protein	not assigned.no ontology
At5g62720	3.10E-06	0.0002776	integral membrane HPP family protein	not assigned.no ontology
At2g15890	3.47E-07	8.22E-05	expressed protein	not assigned.unknown
At1g74440	2.81E-06	0.0002686	expressed protein	not assigned.no ontology
At3g04550	3.43E-07	8.22E-05	expressed protein	not assigned.unknown
At1g16840	1.98E-06	0.0002134	expressed protein	not assigned.unknown
At4g02370	3.57E-06	0.0003055	expressed protein	not assigned.unknown
At4g32340	1.38E-07	4.62E-05	expressed protein	not assigned.unknown
At3g47160	1.76E-07	5.48E-05	expressed protein	not assigned.no ontology
At3g61100	6.64E-07	0.0001111	expressed protein	not assigned.unknown
At5g54970	5.70E-07	0.0001053	expressed protein	not assigned.unknown

Cluster 7

Locus ID	p-value	q-value	Annotation	MapMan Category
At2g37680	4.23E-07	9.16E-05	phytochrome A specific signal transduction component	signalling.light
At3g20810	2.33E-08	2.34E-05	transcription factor jumonji (jmjC) domain-containing protein	RNA.regulation of transcription.JUMONJI family
At3g48360	1.58E-07	5.13E-05	speckle-type POZ protein	RNA.regulation of transcription.Transcriptional Adaptor Zinc Bundle (TAZ) domain family
At3g60530	1.79E-06	0.0002046	zinc finger (GATA type) family protein	RNA.regulation of transcription.C2C2(Zn) GATA transcription factor family
At4g05330	5.84E-07	0.0001053	zinc finger and C2 domain protein	RNA.regulation of transcription.C2H2 zinc finger family
At4g36970	4.99E-07	9.78E-05	remorin family protein	RNA.regulation of transcription.unclassified
At5g03470	1.87E-06	0.0002095	serine/threonine protein phosphatase 2A	protein.postranslational modification
At3g59310	1.62E-06	0.0001919	expressed protein	protein.targeting
At2g38290	2.92E-06	0.0002741	ammonium transporter 2	transport.ammonium
At2g02390	2.87E-06	0.0002718	glutathione S-transferase zeta 1	misc.glutathione S transferases
At1g47210	3.63E-06	0.0003058	cyclin family protein	cell.cycle
At1g28330	5.88E-07	0.0001053	dormancy-associated protein	development.unspecified
At4g21450	2.24E-07	6.37E-05	vesicle-associated membrane family protein	not assigned.no ontology
At3g49780	5.34E-07	0.0001021	phytosulfokines 3	not assigned.no ontology
At5g10860	6.61E-07	0.0001111	CBS domain-containing protein	not assigned.no ontology
At3g18920	3.85E-07	8.76E-05	unknown	
At5g42900	2.24E-08	2.34E-05	expressed protein	not assigned.unknown
At4g30500	3.63E-06	0.0003058	expressed protein	not assigned.unknown
At2g37480	1.93E-06	0.0002131	expressed protein	not assigned.unknown
At1g26665	3.44E-08	2.61E-05	expressed protein	not assigned.unknown
At3g15630	5.10E-07	9.91E-05	expressed protein	not assigned.unknown
At4g19390	9.34E-08	4.17E-05	expressed protein	not assigned.unknown
At3g53670	1.38E-07	4.62E-05	expressed protein	not assigned.unknown
At5g11070	1.32E-06	0.0001668	expressed protein	not assigned.unknown

Table S2. Genes with most significantly changed transcripts in the "Genotype" comparison between antisense-*ACLA* and WT over the SD diurnal cycle (FDR is controlled at the level of 0.2). Supplement to Figure 10A. Carbohydrate metabolic genes, genes involved in transcription.

Cluster 1

Locus ID	p-value	q-value	Annotation	MapMan Category
At4g14090	0.00011	0.1235	anthocyanidin 5-O-glucosyltransferase	misc.UDP glucosyl and glucoronyl transferases
At1g61810	3.51E-05	0.10362	glycosyl hydrolase family 1 protein	misc.gluco-, galacto- and mannosidases
At4g19840	0.00055	0.1891	lectin-related	misc.myrosinases-lectin-jacalin
At4g17030	0.00019	0.13796	expansin-related	cell wall.modification
At2g45400	6.49E-05	0.12212	dihydroflavonol 4-reductase family	secondary metabolism.flavonoids.flavonols
At2g03750	9.21E-05	0.1235	sulfotransferase family protein	lipid metabolism.'exotics' (steroids, squalene etc)
At2g37870	1.39E-06	0.03134	protease inhibitor/seed storage/lipid transfer protein (LTP)	misc.protease inhibitor/seed storage/lipid transfer protein (LTP) family protein
At2g28190	0.00011	0.1235	superoxide dismutase	redox.dismutases and catalases
At1g68570	0.00055	0.1891	proton-dependent oligopeptide transport	transport.peptides and oligopeptides
At1g17350	0.0006	0.19786	auxin-induced protein	hormone metabolism.auxin.induced-regulated-responsive-activated
At4g37580	3.69E-05	0.10362	N-acetyltransferase	hormone metabolism.ethylene.synthesis-degradation
At4g15910	0.00014	0.13796	drought-responsive protei	stress.abiotic.drought/salt
At4g23670	0.00051	0.1891	major latex protein-related	stress.abiotic.unspecified
At2g06190	0.0003	0.15473	gypsy-like retrotransposon family	DNA.synthesis/chromatin structure
At4g37510	0.00019	0.13796	ribonuclease III family protein	RNA.processing
At3g25710	0.00049	0.18888	basic helix-loop-helix (bHLH) family protein	RNA.regulation of transcription.bHLH,Basic Helix-Loop-Helix family
At1g75240	0.00029	0.15376	zinc finger homeobox family protein	RNA.regulation of transcription.unclassified
At3g06160	1.27E-05	0.07158	transcriptional factor B3 family protein	RNA.regulation of transcription.B3 transcription factor family
At5g35330	0.0002	0.13796	methyl-CpG-binding domain-containing protein	RNA.regulation of transcription.Methyl binding domain proteins
At1g76010	0.0004	0.17154	expressed protein	RNA.regulation of transcription.putative DNA-binding protein
At3g16560	0.00039	0.17154	protein phosphatase 2C-relate	protein.postranslational modification
At4g07670	0.00021	0.13796	protease-associated (PA) domain-containing protein	protein.degradation
At5g64150	0.00041	0.17586	methylase family protein	not assigned.no ontology
At2g40490	9.30E-06	0.07158	uroporphyrinogen decarboxylase	tetrapyrrole synthesis
At3g50440	0.00023	0.13796	similar to esterase	not assigned.no ontology
At2g46640	0.00036	0.17154	expressed protein	not assigned.unknown
At1g63420	0.00039	0.17154	expressed protein	not assigned.unknown
At3g59670	0.00025	0.14249	expressed protein	not assigned.unknown
At5g02550	0.0005	0.18888	expressed protein	not assigned.unknown

Cluster 2

Locus ID	p-value	q-value	Annotation	MapMan Category
At4g34135	0.00018	0.13796	UDP-glucoronosyl/UDP-glucosyl transferase	misc.UDP glucosyl and glucoronyl transferases
At1g59960	4.76E-05	0.10362	aldo/keto reductase	minor CHO metabolism.others
At1g35230	0.00044	0.18107	arabinogalactan-protein	cell wall.cell wall proteins.AGPs
At5g05600	0.00017	0.13796	flavonol synthase	secondary metabolism.flavonoids.anthocyanins

At5g35735	0.00038	0.17154	auxin-responsive family protein	hormone metabolism.auxin.induced-regulated-responsive-activated
At5g20810	0.00047	0.18543	auxin-responsive protein	hormone metabolism.auxin.induced-regulated-responsive-activated
At1g75030	0.00024	0.13982	pathogenesis-related thaumatin family protein	stress.biotic
At3g06050	0.00055	0.1891	redox homeostasis	stress.abiotic.unspecified; redox.periredoxins
At3g25010	0.00055	0.1891	disease resistance family protein	stress.biotic
At2g26020	0.00021	0.13796	plant defensin-fusion protein	metal handling.binding, chelation and storage; stress.biotic
At5g44420	0.00011	0.1235	plant defensin protein	metal handling.binding, chelation and storage; stress.biotic
At4g04150	0.00017	0.13796	gypsy-like retrotransposon family	DNA.synthesis/chromatin structure.retrotransposon/transposase.gypsy-like retrotransposon
At1g80840	0.00028	0.15376	WRKY family transcription factor	RNA.regulation of transcription.WRKY domain transcription factor family
At3g57230	0.00023	0.13796	Root Expressed; MADS-box protein	RNA.regulation of transcription.MADS box transcription factor family
At5g03720	9.69E-05	0.1235	member of Heat Stress Transcription Facto	stress.abiotic.heat; RNA.regulation of transcription.HSF,Heat-shock transcription factor family
At1g25220	0.00036	0.17154	Catalyzes the first step of tryptophan biosynthesis	amino acid metabolism.synthesis.aromatic aa.tryptophan
At1g76360	0.0001	0.1235	protein kinase	protein.postranslational modification
At5g52120	0.00046	0.18369	F-box family protein	protein.degradation.ubiquitin.E3.SCF.FBOX; cell.organisation
At3g01920	0.0001	0.1235	yrdC family protein	not assigned.no ontology
At3g25070	0.00016	0.13796	RPM1-interacting protein 4	not assigned.no ontology
At2g29460	1.63E-05	0.07366	glutathione S-transferase	misc.glutathione S transferases
At1g15140	8.50E-05	0.1235	oxidoreductase NAD-binding domain-containing protein	not assigned.no ontology
At5g52760	0.00011	0.1235	heavy-metal-associated domain-containing protein	not assigned.no ontology
At5g52750	0.0005	0.18888	heavy-metal-associated domain-containing protein	not assigned.no ontology
At1g65400	0.0003	0.15473	unknown	
At4g07380	5.05E-05	0.10362	expressed protein	not assigned.unknown
At4g29780	0.00018	0.13796	expressed protein	not assigned.unknown
At5g08240	0.00059	0.19743	expressed protein	not assigned.unknown

Cluster 3

Locus ID	p-value	q-value	Annotation	MapMan Category
At3g29810	0.00037	0.17154	phytochelatin synthetase family protein	cell wall.cellulose synthesis
At4g38880	3.05E-05	0.10362	amidophosphoribosyltransferase	nucleotide metabolism.synthesis.purine
At3g19950	0.00057	0.19297	zinc finger (C3HC4-type RING finger) family protein	RNA.regulation of transcription.unclassified
At1g72130	0.0002	0.13796	proton-dependent oligopeptide transport	transport.peptides and oligopeptides
At5g62730	0.0002	0.13796	proton-dependent oligopeptide transport	transport.peptides and oligopeptides
At3g26320	0.00043	0.18045	cytochrome P450 71B36	misc.cytochrome P450
At5g67310	4.33E-05	0.10362	cytochrome P450 family protein	misc.cytochrome P450
At3g17920	0.00034	0.17099	leucine-rich repeat family protein	not assigned.no ontology
At1g26360	0.00023	0.13796	a six-member gene family	not assigned.no ontology
At1g67330	0.00021	0.13796	expressed protein	not assigned.unknown
At1g13970	1.20E-05	0.07158	expressed protein	not assigned.unknown
At5g54300	0.00054	0.1891	expressed protein	not assigned.unknown

Table S3. Genes with most significantly changed transcripts in the "Genotype*Time" comparison between antisense-*ACLA* and WT over the SD diurnal cycle (FDR is controlled at the level of 0.45). Supplement to Figure 10B. Carbohydrate metabolic genes, genes involved in transcription.

Cluster 1

Locus ID	p-value	q-value	Annotation	MapMan Category
At1g26480	0.0011579	0.4190126	14-3-3 protein GF14 iota (GRF12)	signalling.14-3-3 proteins
At2g15480	0.0006609	0.4190126	UDP-glucoronosyl/UDP-glucosyl transferase family protein	misc.UDP glucosyl and glucoronyl transferases
At3g54920	0.0014628	0.4190126	pectate lyase	cell wall.degradation.pectate lyases and polygalacturonases
At4g20330	0.0015287	0.4190126	transcription initiation factor-related	RNA.processing
At2g18030	0.0002338	0.306494	peptide methionine sulfoxide reductase family protein	protein.aa activation
At1g65670	0.0015158	0.4190126	cytochrome P450 family protein	misc.cytochrome P450
At3g59670	0.0018365	0.4190126	expressed protein	not assigned.unknown
At5g10210	0.0002368	0.306494	expressed protein	not assigned.unknown
At5g55710	0.001277	0.4190126	expressed protein	not assigned.unknown
At4g07380	3.02E-05	0.1238788	expressed protein	not assigned.unknown
At1g05430	0.0008976	0.4190126	expressed protein	not assigned.unknown
At2g42130	0.001016	0.4190126	expressed protein	not assigned.unknown

Cluster 2

Locus ID	p-value	q-value	Annotation	MapMan Category
At2g44480	0.0008899	0.4190126	glycosyl hydrolase family 1	misc.gluco-, galacto- and mannosidases
At4g01890	0.0011546	0.4190126	glycoside hydrolase family 28 protein	cell wall.degradation.pectate lyases and polygalacturonases
At1g61290	0.0013614	0.4190126	syntaxin	cell. vesicle transport
At1g11040	0.0017662	0.4190126	DNAJ chaperone C-terminal domain-containing protein	stress.abiotic.heat
At4g14640	0.0005617	0.4190126	calmodulin-8 (CAM8)	signalling.calcium
At3g24220	0.0016779	0.4190126	9-cis-epoxycarotenoid dioxygenase	hormone metabolism.abscisic acid.synthesis-degradation
At4g07590	0.000201	0.306494	gypsy-like retrotransposon family	DNA.synthesis/chromatin structure.retrotransposon/transposase.gypsy-like retrotransposon
At5g43800	0.001201	0.4190126	copia-like retrotransposon family	DNA.synthesis/chromatin structure.retrotransposon/transposase.copia-like retrotransposon
At5g05510	0.0006095	0.4190126	protein kinase-related	protein.postranslational modification
At5g37450	0.001845	0.4190126	leucine-rich repeat transmembrane protein kinase	protein.postranslational modification
At1g01600	0.0003247	0.3502616	cytochrome P450	misc.cytochrome P450
At3g62280	0.0013627	0.4190126	GDSL-motif lipase/hydrolase family protein	misc.GDSL-motif lipase
At4g01410	0.0021122	0.4482418	harpin-induced family protein	not assigned.no ontology
At1g33340	0.0015072	0.4190126	epsin N-terminal homology (ENTH) domain-containing protein	not assigned.no ontology.epsin N-terminal homology (ENTH) domain-containing protein
AtMg01050	0.0009043	0.4190126	hypothetical protein	not assigned.no ontology
At1g13970	3.83E-05	0.1238788	expressed protein	not assigned.unknown
At1g01690	0.0009619	0.4190126	expressed protein	not assigned.unknown
At5g54790	0.0018046	0.4190126	expressed protein	not assigned.unknown

At4g39190	0.0016996	0.4190126	expressed protein	not assigned.unknown

Cluster 3

Locus ID	p-value	q-value	Annotation	MapMan Category
At1g47960	0.001194	0.4190126	invertase/pectin methylesterase inhibitor family protein	cell wall.pectin*esterases.PME; misc.invertase/pectin methylesterase inhibitor family protein
At4g14090	0.0014615	0.4190126	anthocyanidin 5-O-glucosyltransferase	misc.UDP glucosyl and glucoronyl transferases
At1g61810	0.000441	0.3806333	glycosyl hydrolase family 1 protein	misc.gluco-, galacto- and mannosidases
At3g18080	0.0006415	0.4190126	glycosyl hydrolase family 1 protein	misc.gluco-, galacto- and mannosidases
At2g05260	0.002081	0.4482418	lipase class 3 family protein	lipid metabolism.lipid degradation.lipases
At3g06160	0.0013065	0.4190126	transcriptional factor B3 family protein	RNA.regulation of transcription.B3 transcription factor family
At2g40490	0.0007156	0.4190126	uroporphyrinogen decarboxylase	tetrapyrrole synthesis
At1g63150	2.90E-05	0.1238788	pentatricopeptide (PPR) repeat-containing protein	not assigned.no ontology.pentatricopeptide (PPR) repeat-containing protein
At4g02240	0.002028	0.4482418	unknown	
At1g53620	2.98E-05	0.1238788	hypothetical protein	not assigned.no ontology.glycine rich proteins
At2g02570	0.0013419	0.4190126	expressed protein	not assigned.unknown
At2g16990	0.0017912	0.4190126	expressed protein	not assigned.unknown
At4g11870	0.002081	0.4482418	expressed protein	not assigned.unknown

Cluster 4

Locus ID	p-value	q-value	Annotation	MapMan Category
At1g62980	0.001313	0.4190126	expansin	cell wall.modification
At5g08030	0.0004111	0.3806333	glycerophosphoryl diester phosphodiesterase family protein	lipid metabolism.lipid degradation.lysophospholipases
At1g72890	6.57E-05	0.1700352	disease resistance protein	stress.biotic
At4g37580	0.001811	0.4190126	N-acetyltransferase	hormone metabolism.ethylene.synthesis-degradation
At2g03150	0.0011579	0.4190126	ATP/GTP-binding protein	signalling.G-proteins
At3g31340	0.0001483	0.306494	gypsy-like retrotransposon family	DNA.synthesis/chromatin structure.retrotransposon/transposase.gypsy-like retrotransposon
At2g19640	0.0004262	0.3806333	SET domain-containing protein	RNA.regulation of transcription.SET-domain transcriptional regulator family
At1g76360	0.0006505	0.4190126	protein kinase	protein.postranslational modification
At3g18120	0.0015143	0.4190126	F-box family protein-related	protein.degradation.ubiquitin.E3.SCF.FBOX
At3g63120	0.0016919	0.4190126	cyclin family protein	cell.cycle
At5g13170	0.000775	0.4190126	nodulin MtN3 family protein	development.unspecified
At5g67310	0.0002954	0.3476497	cytochrome P450 family protein	misc.cytochrome P450
At4g21320	0.0015119	0.4190126	heat-stress-associated 32-kD protein	not assigned.unknown
At1g11710	0.0015843	0.4190126	pentatricopeptide (PPR) repeat-containing protein	not assigned.no ontology.pentatricopeptide (PPR) repeat-containing protein
AtMg00850	0.0002299	0.306494	hypothetical protein	not assigned.unknown
At5g15890	0.0008317	0.4190126	expressed protein	not assigned.unknown
At4g02170	0.0008664	0.4190126	expressed protein	not assigned.unknown

APPENDIX F. SUPPLEMENTAL DATA FOR CHAPTER 4

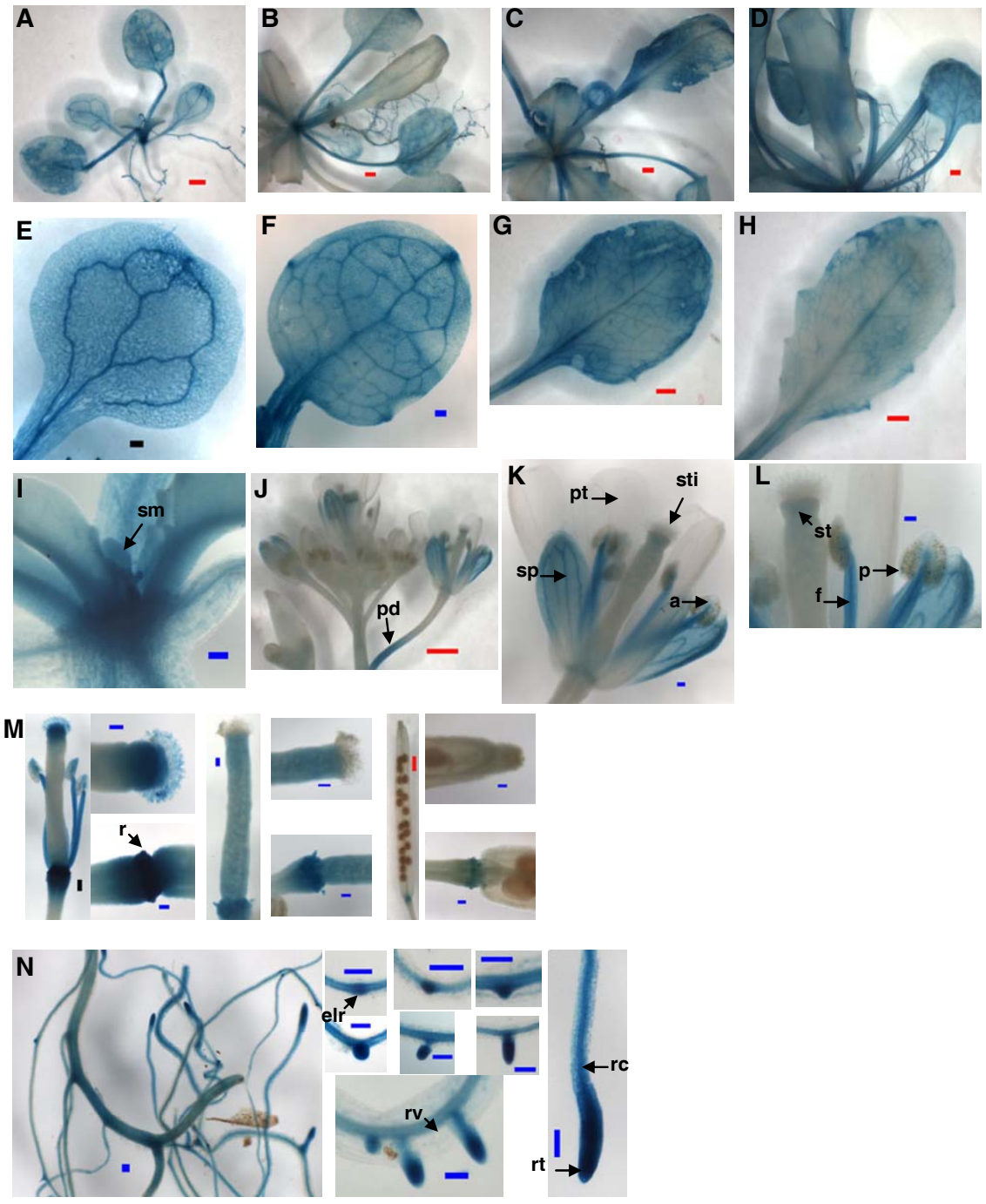


Figure S1.

Figure S1. (continued). *QQ* expression in WT.

QQ expression in seedling (A), plants (B, C, D), cotyledon (E), young leaf (F), old leaf (G, H), shoot meristem (I), flowers and buds (J), flower (K), pollen (L), silique (M), and roots (N). In WT, *QQ* is expressed in young seedling. It is expressed in hypocotyl, cotyledon, leaf mesophyll cell and vasculature. It is expressed in hydathode, trichome and petiole. It is present in shoot meristem. It is highly expressed in roots. It is present in root vasculature, cortex, and root tip. It is expressed in emerging lateral root. As plants grow, it is still expressed in old leaf mesophyll cell and leaf vein. It is expressed in flowers in sepal, filament, pollen, stigma, style, ovary wall and receptacle. It is present in ovule, young silique wall. As silique matures, it is not detectable in silique wall. It is still highly expressed in roots when plants get old. a=anther, elr=emerging lateral root, f=filament, p=pollen, pd=pedicle, pt=petal, r=receptacle, rc=root cortex, rt=root tip, rv=root vasculature, sm=shoot meristem, sp=sepal, sti=stigma, st=style, t=trichome. Red bar=1mm, blue bar=0.1mm.

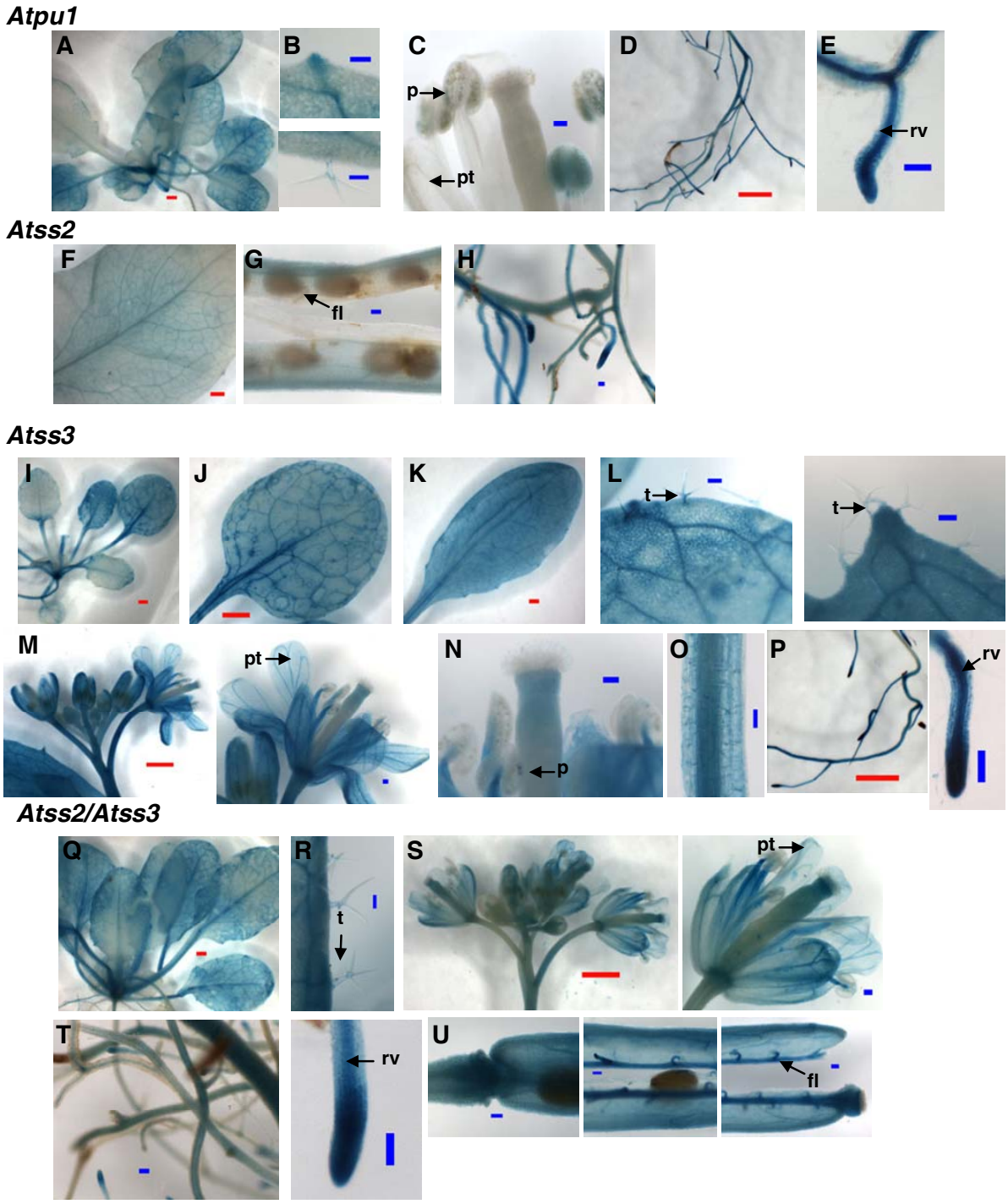

Atpu1
A
B
C
p
pt
D
E
rv
Atss2
F
G
fl
H
Atss3
I
J
K
L
t
t
M
pt
N
p
O
P
rv
Atss2/Atss3
Q
R
t
S
pt
T
rv
U
fl

Figure S2.

Figure S2. (continued). *QQ* expression in the *Atpu1*, *Atss2*, *Atss3*, and *Atss2/Atss3* mutant backgrounds.

QQ expression in plants (A, I, Q), leaves (B, J, K, L, R), flowers (M, S), pollen (C, N), silique (G, U), roots (D, E, H, P, T), and hypocotyl (O). The pattern of expression of QQ in the *Atpu1* mutant is indistinguishable from that of WT. *QQ* has different expression in the starch synthases mutants. In the *Atss3* mutant, *QQ* RNA accumulation is higher than in WT. The *Atss2* mutant has lower *QQ* RNA accumulation than WT (or the *Atss3* mutant). *QQ* RNA accumulation pattern in the *Atss2*/*Atss3* double mutant is similar but slightly reduced compared to the *Atss3* mutant. *QQ* expression is low in the *Atss2* mutant except the high expression in roots. It is expressed low in leaf mesophyll cell, vasculature and silique wall. It is not expressed in funiculus in the silique. In the *Atss3* and *Atss2*/*Atss3* mutants, including the similar but higher expression in leaf (leaf mesophyll cell, leaf vasculature, hydathode, trichome), *QQ* is detectable in petal; it is also highly expressed in silique wall and in funiculus. These are the big differences between the *Atss3*, *Atss2*/*Atss3* mutants and WT, the *Atpu1* and *Atss2* mutants. fl= funiculus, p=pollen, pt=petal, rv=root vasculature, t=trichome. Red bar=1mm, blue bar=0.1mm.

Table S1. Genes with most significantly changed transcripts for the "Genotype" comparison between *ss3* and WT over the SD diurnal cycle. 175 genes are significant when the FDR is controlled at the 0.162 level. The 175 genes are grouped into 5 clusters based on their expression patterns. The gene (At3g30720, named QQ), carbohydrate metabolic genes, genes involved in transcription.

Cluster 1

Locus ID	p-value	q-value	Annotation	MapMan Category
At1g32900	0.00155	0.1618	GBSS	major CHO metabolism.synthesis.starch.starch synthase
At3g46670	0.00105	0.1618	UDP-glucoronosyl/UDP-glucosyl transferase	misc.UDP glucosyl and glucoronyl transferases
At1g02640	0.00012	0.11486	glycosyl hydrolase family 3 protein	cell wall.degradation.mannan-xylose-arabinose-fucose
At1g77410	0.00157	0.1618	beta-galactosidase	cell wall.degradation.mannan-xylose-arabinose-fucose; misc.gluco-, galacto- and mannosidases
At4g15290	0.00109	0.1618	cellulose synthase	cell wall.cellulose synthesis
At3g13190	0.00144	0.1618	myosin heavy chain-related	cell.organisation
At1g16280	0.00084	0.15705	DEAD/DEAH box helicase	DNA.synthesis/chromatin structure
At5g60470	0.00155	0.1618	zinc finger (C2H2 type) family protein	RNA.regulation of transcription.C2H2 zinc finger family
At1g55750	1.47E-05	0.03657	transcription factor-related	RNA.regulation of transcription.unclassified
At1g53690	0.00091	0.15705	DNA-directed RNA polymerases I, II, and III 7 kDa subunit	RNA.transcription
At3g13470	0.0007	0.15705	chaperonin	protein.postranslational modification
At4g24860	0.0017	0.1618	AAA-type ATPase family protein	protein.degradation.AAA type
At3g09790	0.00113	0.1618	polyubiquitin (UBQ8)	protein.degradation.ubiquitin.ubiquitin
At5g09470	6.97E-05	0.10653	mitochondrial substrate carrier family protein	transport.metabolite transporters at the mitochondrial membrane
At1g15500	0.00137	0.1618	chloroplast ADP, ATP carrier protein	transport.metabolite transporters at the mitochondrial membrane
At4g01100	0.00076	0.15705	mitochondrial substrate carrier family protein	transport.metabolite transporters at the mitochondrial membrane
At4g30110	0.00033	0.12979	ATPase E1-E2 type family protein	transport.metal
At4g03400	0.00173	0.1618	auxin-responsive GH3 family protein	hormone metabolism.auxin.induced-regulated-responsive-activated
At5g09590	0.00031	0.12979	heat shock protein 70 / HSP70 (HSC70-5)	stress.abiotic.heat
At5g17000	0.001	0.1618	NADP-dependent oxidoreductase	misc.oxidases - copper, flavone etc.
At4g37400	0.00107	0.1618	cytochrome P450	misc.cytochrome P450
At4g27520	0.00189	0.1618	plastocyanin-like domain-containing protein	misc.plastocyanin-like
At5g04070	0.0003	0.12979	short-chain dehydrogenase/reductase	misc.short chain dehydrogenase/reductase (SDR)
At5g60960	0.00157	0.1618	pentatricopeptide (PPR) repeat-containing protein	not assigned.no ontology.pentatricopeptide (PPR) repeat-containing protein
At2g39710	0.00045	0.14311	aspartyl protease family protein	not assigned.unknown
At1g23600	0.0008	0.15705	expressed protein	not assigned.unknown
At2g04460	0.00124	0.1618	expressed protein	not assigned.unknown
At2g40400	0.00079	0.15705	expressed protein	not assigned.unknown
At1g50400	0.00039	0.13544	expressed protein	not assigned.no ontology
At1g74800	0.00171	0.1618	galactosyltransferase family protein	not assigned.unknown
At3g55000	0.00041	0.13544	tonneau family protein	not assigned.no ontology
At3g56670	0.0001	0.1117	expressed protein	not assigned.unknown

At5g02590	0.00158	0.1618	chloroplast lumen common family protein	not assigned.no ontology
At5g14020	0.00082	0.15705	expressed protein	not assigned.unknown
At5g52230	0.00093	0.15705	expressed protein	not assigned.unknown
At5g52290	0.00122	0.1618	expressed protein	not assigned.unknown
At5g52580	0.00189	0.1618	expressed protein	not assigned.unknown

Cluster 2

Locus ID	p-value	q-value	Annotation	MapMan Category
At2g28950	0.0015	0.1618	expansin	cell wall.modification
At1g14720	0.00187	0.1618	xyloglucan:xyloglucosyl transferase	cell wall.modification
At1g71990	0.00056	0.15502	alpha-(1,4)-fucosyltransferase	cell wall.hemicellulose synthesis
At1g74790	0.00019	0.12979	expressed protein	cell.organisation
At2g36900	0.00127	0.1618	Golgi SNARE protein membrin 11	cell. vesicle transport
At1g75880	0.00087	0.15705	family II extracellular lipase 1	development.unspecified
At1g06850	0.00027	0.12979	bZIP transcription factor	RNA.regulation of transcription.bZIP transcription factor family
At1g49200	0.00081	0.15705	zinc finger (C3HC4-type RING finger)	RNA.regulation of transcription.unclassified
At3g22900	0.0006	0.15705	RNA polymerase Rpb7 N-terminal domain-containing protein	RNA.transcription
At3g12630	0.00105	0.1618	zinc finger (AN1-like) family protein	RNA.regulation of transcription.unclassified
At5g28770	0.00188	0.1618	bZIP transcription factor	RNA.regulation of transcription.bZIP transcription factor family
At3g26900	0.00182	0.1618	shikimate kinase family protein	amino acid metabolism.synthesis.aromatic aa.chorismate
At1g15390	0.00157	0.1618	peptide deformylase	misc.misc2
At3g23880	0.00074	0.15705	F-box family protein	protein.degradation.ubiquitin.E3.SCF.FBOX
At3g62880	0.00087	0.15705	Homologous to pea OEP16 and barley pPORA	transport.protein transport mitochondrial membrane (TIM/TOM)
At1g12840	0.00143	0.1618	vacuolar ATP synthase subunit C	transport.p- and v-ATPases
At1g61100	0.00187	0.1618	disease resistance protein	stress.biotic
At5g54855	0.00163	0.1618	pollen Ole e 1 allergen and extensin family protein	stress.abiotic.unspecified
At5g59530	0.00066	0.15705	2-oxoglutarate-dependent dioxygenase	hormone metabolism.ethylene.synthesis-degradation
At1g07440	0.00176	0.1618	tropinone reductase	secondary metabolism.N misc.alkaloid-like; misc.nitrilases, *nitrile lyases, berberine bridge enzymes, reticuline oxidases, troponine reductases
At1g69740	0.00176	0.1618	porphobilinogen synthase	tetrapyrrole synthesis
At2g02500	0.00108	0.1618	expressed protein	not assigned.unknown
At2g39570	0.00033	0.12979	ACT domain-containing protein	not assigned.no ontology
At2g45980	0.00025	0.12979	expressed protein	not assigned.unknown
At2g29020	0.00064	0.15705	Rab5-interacting family protein	not assigned.no ontology
At2g27830	0.00181	0.1618	expressed protein	not assigned.unknown
At1g22700	0.00169	0.1618	tetratricopeptide repeat (TPR)-containing protein	not assigned.no ontology.pentatricopeptide (PPR) repeat-containing protein
At1g80180	0.00159	0.1618	expressed protein	not assigned.unknown
At1g56500	0.00177	0.1618	haloacid dehalogenase-like hydrolase	not assigned.no ontology
At4g24750	0.00125	0.1618	rhodanese-like domain-containing protein	not assigned.unknown
At4g27750	0.0011	0.1618	expressed protein	not assigned.unknown
At4g29960	9.35E-06	0.02797	expressed protein	not assigned.unknown

At4g35470	0.0004	0.13544	leucine-rich repeat family protein	not assigned.no ontology
At4g39040	0.00062	0.15705	expressed protein	not assigned.unknown
At3g49790	0.0003	0.12979	expressed protein	not assigned.unknown
At5g24610	0.00092	0.15705	expressed protein	not assigned.unknown
At5g39210	0.00034	0.12979	expressed protein	not assigned.unknown
At5g51400	0.00051	0.15042	expressed protein	not assigned.unknown
At5g64816	0.00124	0.1618	expressed protein	not assigned.unknown
At5g21090	0.00081	0.15705	leucine-rich repeat protein	not assigned.no ontology

Cluster 3

Locus ID	p-value	q-value	Annotation	MapMan Category
At5g18670	0.00028	0.12979	BAM9	major CHO metabolism.degradation.starch.starch cleavage
At1g11720	5.25E-06	0.02619	SS3	major CHO metabolism.synthesis.starch.starch synthase
At1g01800	0.00177	0.1618	short-chain dehydrogenase	misc.short chain dehydrogenase/reductase (SDR)
At2g22540	0.00152	0.1618	short vegetative phase protein	RNA.regulation of transcription.MADS box transcription factor family; development.unspecified
At2g38560	0.00093	0.15705	transcription factor S-II	RNA.regulation of transcription.putative transcription regulator; protein.synthesis.elongation
At2g02470	2.02E-05	0.04308	PHD finger family protein	RNA.regulation of transcription.Alfin-like
At3g54360	0.00129	0.1618	expressed protein	RNA.regulation of transcription.putative DNA-binding protein
At5g02840	0.0001	0.1117	myb family transcription factor	RNA.regulation of transcription.MYB-related transcription factor family
At5g47370	0.00095	0.15772	homeobox-leucine zipper protein 2	RNA.regulation of transcription.HB,Homeobox transcription factor family
At3g04940	0.00168	0.1618	cysteine synthase	amino acid metabolism.synthesis.serine-glycine-cysteine group.cysteine.OASTL
At1g20650	0.00026	0.12979	protein kinase family protein	protein.postranslational modification
At1g22050	0.00126	0.1618	ubiquitin family protein	protein.synthesis.misc ribososomal protein; protein.degradation.ubiquitin
At5g58870	0.00151	0.1618	FtsH protease	protein.degradation.metalloprotease
At4g32420	0.00014	0.12855	peptidyl-prolyl cis-trans isomerase cyclophilin-type family protein	cell.cycle
At2g37460	0.00019	0.12979	nodulin MtN21 family protein	development.unspecified
At5g49730	0.0012	0.1618	a putative ferric chelate reductase	metal handling.acquisition
At1g56280	0.0003	0.12979	drought-responsive family protein	stress.abiotic.drought/salt
At3g62190	0.00188	0.1618	DNAJ heat shock N-terminal domain-containing protein	stress.abiotic.heat
At5g38350	0.00027	0.12979	disease resistance protein	stress.biotic
At5g49480	0.00187	0.1618	sodium-inducible calcium-binding protein	signalling.calcium
At2g41410	0.00093	0.15705	calmodulin	signalling.calcium
At1g64850	0.00108	0.1618	calcium-binding EF hand family protein	signalling.calcium
At2g38210	0.00088	0.15705	ethylene-responsive protein	hormone metabolism.ethylene.induced-regulated-responsive-activated
At2g22670	0.00125	0.1618	auxin-responsive protein	hormone metabolism.auxin.induced-regulated-responsive-activated; RNA.regulation of transcription.Aux/IAA family
At3g04730	0.00125	0.1618	auxin-responsive protein	hormone metabolism.auxin.induced-regulated-responsive-activated; RNA.regulation of transcription.Aux/IAA family
At5g47970	0.00016	0.12855	nitrogen regulation family protein	N-metabolism.misc

At5g65480	0.0012	0.1618	expressed protein	not assigned.unknown
At1g10660	0.00084	0.15705	expressed protein	not assigned.unknown
At2g14730	0.0013	0.1618	expressed protein	not assigned.unknown
At1g76020	0.00168	0.1618	expressed protein	not assigned.unknown
At1g55960	0.00088	0.15705	expressed protein	not assigned.no ontology
At1g29530	0.00073	0.15705	expressed protein	not assigned.unknown
At3g21260	0.00143	0.1618	glycolipid transfer protein-related	not assigned.no ontology
At1g33490	0.00022	0.12979	expressed protein	not assigned.unknown
At2g33250	0.00035	0.12996	expressed protein	not assigned.unknown
At4g13840	7.12E-05	0.10653	transferase family protein	not assigned.no ontology
At4g29070	9.25E-06	0.02797	expressed protein	not assigned.unknown

Cluster 4

Locus ID	p-value	q-value	Annotation	MapMan Category
At3g30720	1.56E-06	0.02339	QQ	not assigned.unknown
At1g30620	0.00055	0.15502	UDP-D-xylose 4-epimerase	cell wall.precursor synthesis.MUR4
At3g53520	0.00124	0.1618	NAD-dependent epimerase/dehydratase	cell wall.precursor synthesis.UXS
At1g58370	0.00158	0.1618	glycosyl hydrolase family 10 protein	cell wall.degradation.mannan-xylose-arabinose-fucose
At2g18280	0.00146	0.1618	Member of TLP family	cell.organisation
At3g49120	0.00048	0.14793	Class III peroxidase Perx34	misc.peroxidases
At3g15980	0.00123	0.1618	coatomer protein complex	cell. vesicle transport
At4g14470	0.00071	0.15705	non-LTR retrotransposon family	DNA.synthesis/chromatin structure
At1g60930	0.0011	0.1618	DNA helicase	DNA.unspecified
At1g76880	0.00148	0.1618	trihelix DNA-binding protein	RNA.regulation of transcription.Trihelix, Triple-Helix transcription factor family
At1g68130	0.00024	0.12979	zinc finger (C2H2 type) family protein	RNA.regulation of transcription.C2H2 zinc finger family
At4g37540	0.00034	0.12979	LOB domain protein 39	RNA.regulation of transcription.AS2,Lateral Organ Boundaries Gene Family- Class II
At5g25920	0.00172	0.1618	expressed protein	RNA.regulation of transcription.putative DNA-binding protein
At4g36710	3.09E-05	0.05786	scarecrow transcription factor family protein	RNA.regulation of transcription.GRAS transcription factor family; development.unspecified
At5g65010	0.00131	0.1618	asparagine synthetase	amino acid metabolism.synthesis.aspartate family.asparagine
At3g01350	0.00016	0.12855	proton-dependent oligopeptide transport	transport.peptides and oligopeptides
At4g30820	0.00069	0.15705	cyclin-dependent kinase-activating kinase	protein.postranslational modification; cell.organisation; cell.cycle
At3g23750	0.00082	0.15705	leucine-rich repeat family protein	protein.postranslational modification
At1g75990	0.00147	0.1618	26S proteasome regulatory subunit S3	protein.degradation.ubiquitin.proteasom
At3g22890	0.00117	0.1618	ATP sulfurylase	S-assimilation.APS
At4g01090	0.00169	0.1618	extra-large G-protein-related	signalling.G-proteins
At3g57530	0.00189	0.1618	calcium-dependent protein kinase	signalling.calcium
At5g14040	8.64E-05	0.1117	mitochondrial phosphate transporter	transport.phosphate; transport.metabolite transporters at the mitochondrial membrane
At2g44750	0.00176	0.1618	thiamin pyrophosphokinase	Co-factor and vitamine metabolism
At2g18440	0.00187	0.1618	expressed protein	not assigned.unknown
At1g64570	0.00118	0.1618	expressed protein	not assigned.unknown
At4g18540	0.00079	0.15705	expressed protein	not assigned.unknown
At5g01080	0.00051	0.15042	expressed protein	not assigned.unknown
At5g10010	0.00093	0.15705	expressed protein	not assigned.unknown

At5g40450	0.00124	0.1618	expressed protein	not assigned.unknown
At5g58100	0.00135	0.1618	expressed protein	not assigned.unknown
At3g51770	0.00116	0.1618	similar to tetratricopeptide repeat (TPR)-containing protein	not assigned.no ontology.pentatricopeptide (PPR) repeat-containing protein
At2g11260	0.00184	0.1618		

Cluster 5

Locus ID	p-value	q-value	Annotation	MapMan Category
At4g09020	0.00163	0.1618	ISA3	major CHO metabolism.synthesis.starch.debranching
At4g26530	0.00149	0.1618	fructose-bisphosphate aldolase	PS.calvin cyle.aldolase; glycolysis.aldolase
At3g06360	0.00039	0.13544	arabinogalactan-protein	cell wall.cell wall proteins.AGPs
At4g16590	0.00059	0.15705	glucosyltransferase-related	cell wall.cellulose synthesis
At5g62360	0.00043	0.14055	invertase/pectin methylesterase inhibitor family protein	misc.invertase/pectin methylesterase inhibitor family protein
At3g61060	0.0011	0.1618	F-box family protein / lectin-related	misc.myrosinases-lectin-jacalin
At1g72040	0.00048	0.14793	deoxynucleoside kinase family	nucleotide metabolism.deoxynucleotide metabolism
At3g10270	0.0014	0.1618		DNA.repair
At3g55060	0.00034	0.12979	expressed protein	DNA.synthesis/chromatin structure
At4g25630	0.0003	0.12979	fibrillarin 2 (FIB2)	RNA.processing
At3g44750	0.00184	0.1618	histone deacetylase	RNA.regulation of transcription.HDA
At2g19810	4.86E-06	0.02619	zinc finger (CCCH-type) family protein	RNA.regulation of transcription.C3H zinc finger family
At4g10320	0.00015	0.12855	isoleucyl-tRNA synthetase	protein.aa activation
At5g66880	0.00186	0.1618	a member of SNF1-related protein kinase	protein.postranslational modification
At2g29310	0.00163	0.1618	tropinone reductase	secondary metabolism.N misc.alkaloid-like; misc.nitrilases, *nitrile lyases, berberine bridge enzymes, reticuline oxidases, troponine reductases
At2g46440	0.00083	0.15705	cyclic nucleotide-regulated ion channel	transport.cyclic nucleotide or calcium regulated channels
At4g01370	0.00138	0.1618	mitogen-activated protein kinase	signalling.MAP kinases
At4g08920	0.0017	0.1618	cryptochrome 1 apoprotein (CRY1)	signalling.light
At5g20820	9.87E-05	0.1117	auxin-responsive protein-related	hormone metabolism.auxin.induced-regulated-responsive-activated
At2g18220	0.00054	0.15502	expressed protein	not assigned.unknown
At1g24490	0.0006	0.15705	Homologue of the Alb3/Oxa1/YidC family	not assigned.unknown
At1g67510	0.00019	0.12979	leucine-rich repeat family protein	not assigned.no ontology
At1g59520	0.00068	0.15705	expressed protein	not assigned.unknown
At1g76590	0.00038	0.13544	zinc-binding family protein	not assigned.no ontology
At1g71300	0.00128	0.1618	Vps52/Sac2 family protein	not assigned.no ontology
At5g02010	0.00085	0.15705	a member of KPP-like gene family	not assigned.unknown
At5g11480	0.00093	0.15705	expressed protein	not assigned.unknown
At1g56530	0.00153	0.1618	hydroxyproline-rich glycoprotein family protein	not assigned.no ontology.hydroxyproline rich proteins

Table S2. *QQ* is expressed differently in WT, *Atpu1*, *Atss2*, *Atss3* and *Atss2*/*Atss3* mutants.

Tissue		**WT**	***Atpu1***	***Atss2***	***Atss3***	***Atss2/Atss3***
Root	Tip	+++	+++	+++	+++	+++
	Vasculature	+++	+++	+++	+++	+++
	Epidermis/cortex	+	+	+	+	+
	Root surface	+	+	+	+	+
	Root meristem	+++	+++	+++	+++	+++
Cotyledon	Petiole	++	+		+	+
	Vasculature	++	++	+	+++	++
	Mesophyll	+	+		+	+
	Hydathode	+	+		+	+
Leaf	Shoot meristem	+	+		+	+
	Petiole	+	+		+++	+
	Vasculature	++	++	+	+++	++
	Mesophyll	+	+		+++	++
	Hydathode	+	+		+	+
	Trichome	+	+		+	+
Stem	Vasculature	+	+		+++	+
	Cortex	+	+		+	+
Flower	Pedicle	+	+	+	+++	++
	Receptacle	+	+	+	+++	++
	Sepal	+	+	+	+++	++
	Petal	-	-	-	+	+
	Filament	+	+	+	+	+
	Anther	+	+	-	+	+
	Pollen	+(Mature)	+(Mature)	-	+(Mature)	+(Mature)
	Ovary wall	+	+	-	+	+
	Style	+	+	-	+	+
	Stigma	+	+	-	+	+
Silique	Wall	+(Young)	+(Young)	+(Young)	+++	+++
	Seed/Embryo	+	+	-	+	+

QQ spatial expression in WT and mutants. “+” means there is signal. “-” means no signal is detected. More “+” means stronger signal.

ACKNOWLEDGEMENTS

First, I would like to express my deep appreciation to my major professor, Dr. Eve Syrkin Wurtele, for her inspiration, encouragement and guidance through all these years. I want to thank her sincerely for the opportunity she has given to me, for what I have learned both from her and also from Wurtele lab. Many thanks to her for her patience and insightful advice on both my research and my dissertation, as well as her friendship in my life, and for her suggestion to give the small unique gene such a pretty name (QQ) after my daughter (Qingqing).

I am very grateful to my POS committee members: Dr. Martha James, Dr. Dan Nettleton, Dr. Basil Nikolau, Dr. Martin Spalding, and Dr. Mark Westgate for the knowledge they taught me, and for their kind help, helpful discussions and thoughtful suggestions to my research and my dissertation.

I sincerely appreciate Dr. Alan Myers' help and advice to my starch research.

Many thanks to the past and current members of Wurtele Lab, Dr. Carol Foster, Dr. Hilal Ilarslan, Dr. Huirong Qian, Dr. Beth Winters, Dr. Wiesia Mentzen, Nick Ransom, Hongjian Liang, Jie Li and other members for warm-hearted help, discussions and friendship over these years.

Many thanks to all my co-workers of starch group, soybean group and MetNet group, for their friendship, help and discussions. Thanks to all my friends and those who have helped me.

Lastly, I would like to thank my family for their support. Thanks to my parents and my husband for their understanding and loving. And finally, I would like to thank my daughter, Qingqing, for her loving, and understanding to let me come to work at the time she wants me.