(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
12 July 2012 (12.07.2012)

WIPO | PCT

(10) International Publication Number
WO 2012/094529 A2

(51) International Patent Classification:
*C07K 14/415* (2006.01) *C12N 15/82* (2006.01)

(21) International Application Number:
PCT/US2012/020375

(22) International Filing Date:
5 January 2012 (05.01.2012)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/430,120 5 January 2011 (05.01.2011) US

(71) Applicant *(for all designated States except US)*: **THE CURATORS OF THE UNIVERSITY OF MISSOURI** [US/US]; 340A Bond Life Sciences Center, 1201 East Rollins, Columbia, MO 65211 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **MITCHUM, Melissa, Goellner** [US/US]; 5000 Forum Blvd., Columbia, MO 65203 (US). **KANDOTH, Pramod, Kaitheri** [IN/US]; 1027 Southpark Drive, Apt. 3, Columbia, MO 65201 (US). **YECKEL, Greg** [US/US]; 5829 Itaska Street, St. Louis, MO 63109 (US). **ITHAL, Nagabhushana** [IN/IN]; Ff03, #17, Pronaam Residency, Amruthanagara, Hebbal, Bangalore 560 092, Karnataka (IN).

(74) Agent: **CHEN, X., Peter**; Lathrop & Gage LLP, 4845 Pearl East Circle, Suite 201, Boulder, CO 80301 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

**Published**:

— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

WO 2012/094529 A2

(54) **Title**: GENES IMPLICATED IN RESISTANCE TO SOYBEAN CYST NEMATODE INFECTION AND METHODS OF THEIR USE

(57) **Abstract**: To gain new insights into the mechanism of soybean resistance to the soybean cyst nematode (*Heterodera glycines*), gene expression profiles of developing syncytia in soybean near-isogenic lines (NILs) differing at the at *Rhg1* locus were compared. Gene expression profiling revealed that 1,447 genes were differentially expressed between the line that is resistant to SCN and the line that is sensitive to SCN infection. These results indicate that syncytia induced in the resistant line are undergoing severe oxidative stress and imbalanced endoplasmic reticulum homeostasis, both of which likely contribute to the resistance reaction. Defense-related genes up-regulated within syncytia of the resistant line included those predominantly involved in apoptotic cell death, the plant hypersensitive response, and salicylic acid (SA)-mediated defense signaling; many of these genes were partially suppressed by a virulent soybean cyst nematode population for successful nematode reproduction and development on the resistant line. Expression of these SCN responsive genes may be manipulated to obtain SCN resistant lines.

# GENES IMPLICATED IN RESISTANCE TO SOYBEAN CYST NEMATODE INFECTION AND METHODS OF THEIR USE

## RELATED APPLICATIONS

This application claims benefit of priority to United States Provisional Patent Application Serial No. 61/430,120 filed January 5, 2011, the contents of which are incorporated by reference into this application.

## BACKGROUND

### 1. Field of the Invention

[0001] The present disclosure relates to proteins and genes responsible for plant defense against soybean cyst nematode (SCN) and the regulation of their expression in plant defense against SCN infection.

### 2. Description of the Related Art

[0002] Soybean cyst nematode (SCN; *Heterodera glycines*) is an obligate sedentary endoparasite, and is the most important pathogen for soybean. The average amount of soybean lost to SCN in the United States from 2006 to 2009 was 128.6 million bushels annually, which was valued at $1.286 billion annually (Koenning and Wrather, 2010). The successful invasion of soybean by SCN depends upon SCN's ability to establish a permanent feeding cell (also known as "syncytium") within the roots of soybean. Infective juveniles penetrate into the root and migrate toward the vasculature. Near the vasculature, each juvenile selects a single cell, which is modified to allow for the incorporation of adjacent cells through progressive cell wall dissolution to form a multinucleate syncytium. The nematode derives nutrients from the syncytium for its growth and reproduction.

[0003] The primary management practice for this pathogen is by developing resistant soybean cultivars. Resistant cultivars have been developed by identifying SCN-resistant soybean germplasm from collections of plant introductions (PI) and incorporating the trait through conventional breeding programs. Although several sources of resistance have been identified, only a few PIs have been used in breeding programs due to undesirable traits associated with other resistance sources. The most predominant sources of resistance found in commercially available cultivars are derived from PI 88788, PI 54840 (Peking), and PI 437654. In all resistant cultivars, the infective juveniles are capable of penetrating into roots and can induce the formation of syncytia, but the syncytia become necrotic soon after establishment and the nematodes starve to death. Although necrosis is a common theme, the timing of necrosis and degeneration of syncytia vary among resistant cultivars, depending on the source of resistance

(Acedo *et al.*, 1984). For example in Peking, syncytial collapse is observed as early as 48 hours post infection (Mahalingam and Skorupska, 1996); whereas, the onset of the resistance response is much slower in PI 209332, with degeneration of the syncytia not occurring until 8-10 days post infection ("dpi") (Acedo *et al.*, 1984).

**[0004]** Despite the extensive histological studies documenting the cellular changes associated with degenerating syncytia in soybean (Endo, 1965; Riggs *et al.*, 1973; Acedo *et al.*, 1984), very little is known about the molecular mechanisms underlying this hypersensitive-like resistance response. Research conducted in the past decade has identified a number of quantitative trait loci (QTLs) associated with SCN resistance (reviewed in Concibido *et al.*, 2004) in different PIs that serve as sources of resistance in breeding programs. Among these, two major QTLs are *Rhg1* on soybean chromosome 18 (formerly linkage group G) and *Rhg4* on chromosome 8 (formerly linkage group A2). *Rhg1* exhibits incomplete dominance and contributes to a significant portion of SCN resistance in most PIs tested, including PI 88788, PI 90763, PI 209332, and Peking (Concibido *et al.*, 2004). In addition, *Rhg1* is effective against a broad spectrum of SCN populations. *Rhg4* is dominant and is required for full resistance to certain SCN populations in some (e.g., Peking, PI 437654), but not all (e.g., PI 209332, PI 88788), resistant sources (Brucker *et al.*, 2005).

**[0005]** Microarray analyses have been carried out to study this plant-nematode interaction. Initial studies used whole soybean roots infected with SCN to assess transcriptional changes during a compatible interaction (Khan *et al.*, 2004; Alkharouf *et al.*, 2006; Ithal *et al.*, 2007a; Klink *et al.*, 2007a). However, due to the specialized nature of the interaction and the location of syncytia well within the root, it is very difficult to draw meaningful conclusions using whole roots to understand this pathosystem.

**[0006]** Laser capture microdissection (LCM) of syncytial cells coupled with microarray analysis has been particularly useful in extending our understanding of the SCN-soybean interaction, as indicated by recently published studies (Klink *et al.*, 2005; Ithal *et al.*, 2007b). These studies have provided new insights into the underlying molecular events occurring during syncytium development. More recently, the same technology has been applied to study incompatible SCN-soybean interactions (Klink *et al.*, 2007b; Klink *et al.*, 2009; Klink *et al.*, 2010). Two studies reported on a comparative microarray analysis of soybean genes induced in response to either a virulent or an avirulent SCN population on Peking (Klink *et al.*, 2007b; Klink *et al.*, 2009), demonstrating that soybean can differentiate between nematode populations prior to feeding cell establishment (Klink *et al.*, 2007b). The same group also published a microarray study that examined the transcriptional changes occurring in syncytia

induced by an avirulent SCN population on PI 88788 at three time points after infection (Klink *et al.*, 2010).

**[0007]** There have been no reports of a direct comparative analysis of syncytia gene expression profiles using near-isogenic lines (NILs) to identify transcripts regulated by specific soybean resistance genes. NILs have several advantages over PIs for comparative analyses of plant gene expression between resistant and susceptible soybean in response to SCN. Theoretically, NILs can share up to 98% of their genome, differing only in a region encompassing a trait of interest (Li *et al.*, 2004); thus, NILs are powerful tools to study the effects of specific gene loci with reduced genetic background effects. Consequently, the use of NILs for molecular studies is becoming more popular. For instance, NILs have been used in a microarray analysis of iron efficient and inefficient cultivars of soybean (O'Rourke *et al.*, 2009) and a wheat leaf rust resistance gene *Lr10* (Manickavelu *et al.*, 2010). NILs have also been used recently to help identify the effects of the Arabidopsis gene *FLC* on seed germination (Chiang *et al.*, 2009). Despite intensive cytological and molecular genetic studies, the genes responsible for SCN resistance have not been identified (Melito *et al.*, 2010), and the mechanism for resistance on a molecular level has yet to be fully elucidated.

## SUMMARY

**[0008]** The instrumentalities described herein overcome the problems outlined above and advance the art by providing genes and DNA regulatory elements in plant roots that may play an important role in plant defense against SCN infection. More specifically, gene expression levels in root tissues are compared between soybean lines that are resistant to SCN infection ("resistant lines) and lines that are susceptible to SCN infection ("susceptible lines"). The expression levels of a number of genes have been found to be significantly higher in the SCN resistant lines as compared to the expression levels of the same genes in the susceptible lines. Conversely, a number of genes have been found to be expressed at a significantly lower level in the SCN resistant lines as compared to the expression levels of the same genes in the susceptible lines. For purpose of this disclosure, all these genes whose levels are up- or down-regulated in the SCN resistant lines as compared to the expression levels of the same genes in the susceptible lines are termed "SCN responsive gene(s)," or "SCNRG." For purpose of this disclosure, a gene having a sequence that is substantially identical to an SCNRG is considered a member of the SCNRG family. A gene encoding a protein having an amino acid sequence that is substantially identical to the amino acid sequence of a protein encoded by an SCNRG is also considered a member of the SCNRG family. In another embodiment, a gene from a different

plant species or from the same species, which shares similar domain structure with a SCNRG disclosed herein and which performs similar cellular function as that of the SCNRG disclosed herein is also considered a member of the SCNRG family, and may be referred to as a homolog of the SCNRG for purpose of this disclosure.

**[0009]** Planting of resistant soybean cultivars has been the primary strategy for managing SCN population levels in the field. Despite widespread use of SCN-resistant soybean, SCN still causes an estimated $1.286 billion annually in yield losses. Lack of understanding of the molecular basis of resistance to this pathogen continues to hinder progress to enhance the effectiveness and durability of natural plant resistance. This lack of knowledge also hinders efforts to design novel strategies for resistance through biotechnological approaches. *Rhg1*, a major resistance locus in almost all SCN-resistant germplasm, is required for resistance against multiple SCN Hg types (Concibido *et al.*, 2004); however, the molecular nature of the resistance gene underlying *Rhg1* (Melito *et al.*, 2010) is not known. Moreover, the downstream signaling and response genes mediated by *Rhg1* have yet to be identified. The *Rhg1* gene has been mapped to chromosome 18 and is within 0.4 cm of SSR marker satt_309 (Cregan *et al.*, 1999), enabling the generation of NILs differing only at this locus (Mudge, 1999). Due to the multigenic nature of SCN resistance, these NILs are very useful for dissecting the SCN-soybean incompatible interaction.

**[0010]** In order to better understand the molecular events associated with *Rhg1*-mediated resistance against SCN, NILs have been used for a comparative analysis of syncytial gene expression using LCM and microarrays. These NILs have been used previously to study the effects of *Rhg1* on root penetration and development of SCN (Li *et al.*, 2004). Although root penetration by SCN juveniles is similar between NIL-R and NIL-S, the growth, development, and fecundity of nematode females is suppressed on NIL-R (Li *et al.*, 2004), suggesting that *Rhg1* may have a negative impact on syncytium development and maintenance. The histological characteristics of syncytia in resistant soybean cultivars to SCN infection are well documented (Ross, 1958; Endo, 1965; Riggs *et al.*, 1973; Acedo *et al.*, 1984). Second-stage SCN juveniles (J2s) induce the formation of syncytia in all resistant cultivars, but syncytial collapse several days later leads to nematode starvation and death. In one type of resistance (PI 437654; Peking type), syncytial collapse is very rapid and begins to occur within 48 hours of induction of the syncytium (Mahalingam and Skorupska, 1996).

**[0011]** To gain new insight into the cause of the aberrant syncytia development that occurs in resistant soybean in response to SCN infection, gene expression is analyzed in syncytia induced in soybean NILs differing at the *Rhg1* locus (Mudge, 1999). Laser capture

microdissection (LCM) coupled with comparative microarray profiling of syncytia isolated from the resistant NIL (NIL-R) and susceptible NIL (NIL-S) results in the identification of 1,447 differentially expressed genes with the false discovery rate set at 10%. Many of the genes induced in the NIL-R are soybean homologs of genes known to play important roles in disease resistance responses of other plant species to various pathogens, including canonical resistance genes (e.g., CC-NB-LRR class of receptors), genes associated with the hypersensitive-like response (HR), apoptotic cell death, the salicylic acid (SA)-mediated resistance pathway, and several transcription factors with defense-related roles. These results have also been validated by syncytia-specific quantitative PCR (qPCR), time course qPCR on infected whole root pieces, and promoter-GUS reporter experiments. These results suggest that *Rhg1* may mediate a complex defense response within syncytia formed in resistant soybean plants, ultimately limiting the growth and development of the nematode.

**[0012]** Because the NILs used in the present disclosure show a delayed type of resistance (Acedo *et al.*, 1984; Li *et al.*, 2004) with notable histological changes to syncytia occurring by 8-10 dpi (Figure 1), 5 and 8 dpi time points are chosen for laser-capture of syncytia to reflect gene expression prior to the onset of syncytium collapse. The comparison of syncytia gene expression between NIL-R and NIL-S by microarray analysis identified 1,447 differentially expressed probe sets (Tables 1-11). A relatively high representation of stress- and defense-related genes have been identified (241 probe sets representing 16.8% of the total; Figure 2; Table 8), including genes involved in oxidative, heat, cold, salt, and drought stress.

**[0013]** In one embodiment, a gene coding for a BAG (BCL2-associated athanogene) domain protein with highest homology to the Arabidopsis BAG6 protein (AtBAG6) is the most highly up-regulated gene in syncytia of the resistant line. BAG proteins are anti-apoptotic in animals; however, *AtBAG6* causes programmed cell death in yeast and Arabidopsis in overexpression studies (Kang *et al.*, 2006). AtBAG6 is up-regulated by heat stress, and the HSF A-2 (Probe set GmaAffx.71308.2.A1_at, 4.0 fold) is involved in its regulation (Nishizawa *et al.*, 2006). Increased expression of BAG6 gene in the resistant line suggests that the syncytia may be undergoing an apoptotic-like cell death response.

**[0014]** Several genes related to ER stress were also found to be up-regulated (e.g., BZIP60 homolog, BIP2 homolog, Calnexin, Bax inhibitor genes, and several protein disulphide isomerases). ER stress is a cellular condition in which unfolded proteins accumulate in the ER. Mis-folding of proteins may be the result of mutations, disturbances in calcium homeostasis, and the heightened need for protein folding. In order to maintain ER homeostasis under such conditions, signaling pathways are activated that are collectively known as the UPR. When ER

stress is not relieved, apoptotic cell death may occur (Urade, 2009). It has recently been reported that water deficit or drought leads to programmed cell death mediated by the ER stress response pathway in Arabidopsis roots (Duan *et al.*, 2010). Several drought and ABA-induced genes were found to be up-regulated in syncytia of the resistant NIL. Taken together, these data suggest that multiple stresses are induced by an upstream signaling event that ultimately leads to the activation of an HR-like programmed cell death (PCD) causing the pathogen to starve and die. It is also possible that pathogen death occurs before the syncytial HR-PCD. The HR may represent the final stages of the resistance response where a certain threshold of defense-related responses has been reached (Morel and Dangl, 1997). For example, the Arabidopsis *dnd1* mutant expresses resistance to pathogens that otherwise induce HR on wild-type plants (Clough *et al.*, 2000).

**[0015]** In general, the SA pathway has been shown to be activated in resistance against biotrophs and the JA pathway has been shown to be activated in resistance to necrotrophs and insects (Glazebrook, 2005; Bari and Jones, 2009). The SA pathway has also been implicated in resistance to the root-knot nematode in tomato (Branch *et al.*, 2004). Here, several homologs of genes belonging to the SA-mediated defense signaling pathway have been identified to be up-regulated in SCN-induced syncytia of the NIL-R lines. These included soybean homologs of Arabidopsis *NDR1* and *NDR1/HIN1*-like (*NHLs*) genes, which are key signal transducers in SA-mediated signaling. NDR1 is a plasma membrane localized protein required for disease resistance to *P. syringae* pv. tomato DC3000 carrying avirulence genes *avrRpm1*, *avrRpt2*, *avrPph3*, and *avrB*. It is also required for resistance against avirulent isolates of the fungal pathogen *Peronospora parasitica* (Century *et al.*, 1995; Century *et al.*, 1997). The requirement for resistance against a diverse group of pathogens suggests that this is a common downstream element in *R*-gene-mediated resistance in plants. Arabidopsis *ndr1* mutants have reduced ROS production and SA accumulation in response to avirulent bacteria (Shapiro and Zhang, 2001). Conversely, overexpression of *NDR1* in Arabidopsis leads to enhanced resistance to virulent *P. syringae* pv. tomato (Coppinger *et al.*, 2004). NHL proteins have sequence homology to NDR1 of Arabidopsis and HIN1 of tobacco and are pathogen-induced in Arabidopsis (Varet *et al.*, 2002). *NHL3* overexpression in Arabidopsis is associated with enhanced resistance to virulent strains of *P. syringae* (Varet *et al.*, 2003). A homolog of Arabidopsis PBS3 (WIN3) has also been identified, which interacts with the *P. syringae* effector protein HopW1-1 and is important for responses induced by several effectors in Arabidopsis. PBS3 is an important component of NDR1-dependent RPS2-mediated resistance against *P. syringae* pv. tomato carrying *avrRpt2* and also plays a role in basal resistance (Lee *et al.*, 2007).

PBS3 is an acyl adenylase, and the Arabidopsis *pbs3* mutant exhibits enhanced susceptibility to *P. syringae* pv. tomato carrying *avrPphB* (Nobuta *et al.*, 2007). In the *pbs3* mutant, induced free and conjugated SA levels are reduced. A homolog of another Arabidopsis gene related to SA accumulation, *WIN1*, is down-regulated in syncytia of NIL-R. Overexpression of *WIN1* delays SA accumulation in response to several effectors, including HopW1-1 (Lee *et al.*, 2008), which indicates WIN1 is a negative regulator. Thus, down-regulation of this gene would be predicted to have a positive impact on SA levels.

**[0016]** It has been previously reported that the majority of JA pathway components are suppressed during a compatible soybean-SCN interaction (Ithal *et al.*, 2007b). Here, a homolog of Arabidopsis lipoxygenase 1 (*AtLOX1*) has been found to be up-regulated in syncytia of the NIL-R, suggesting the involvement of lipid peroxides in the resistance response. The up-regulation of soybean *LOX* genes is also reported in syncytia induced on Peking and PI 88788 by an avirulent population of SCN (Klink *et al.*, 2009; Klink *et al.*, 2010). Lipoxygenases have a role in basal resistance to the root-knot nematode in maize (Gao *et al.*, 2008). Recently, mutations in *AtLOX1* and silencing of a homologous gene of *Capsicum annuum* (*CaLOX1*) have been shown to increase susceptibility to diverse microbial pathogens (Hwang and Hwang, 2010). *CaLOX1*-silenced plants show lowered SA and ROS levels. However, we also identified down-regulation of two soybean genes corresponding to homologs of allene oxide cyclases (AOCs) involved in JA biosynthesis and a homolog of JAR1, a protein required to convert JA to the biologically active JA-isoleucine. These discrepancies emphasize the need for further studies directed at silencing the genes involved in SA and JA biosynthesis and quantifying hormone levels in nematode-infected roots to clarify the role of these small molecules in SCN-induced resistance in soybean.

**[0017]** In another aspect, evidence is presented for the potential involvement of a complex stress and defense-related response, including increased expression of genes involved in the production of ROS, the unfolded protein response, SA-mediated signaling, and plant PCD in *Rhg1*-mediated resistance to SCN. Involvement of almost all hormones shows an intricate network of cross-talk associated with this defense response. The instant disclosure also highlights the importance of conducting a direct comparison between syncytia transcriptomes in the resistant versus susceptible NILs using the same nematode population to identify genes potentially involved in resistance. Inadvertently, a large number of genes would be overlooked in a direct comparison of syncytia transcriptomes induced in the resistant line by an avirulent versus a virulent SCN population. It is shown here that the plant still mounts a defense response against the virulent nematode population, albeit somewhat attenuated compared to its response

against the avirulent nematode population. In contrast, the response of the susceptible line to the avirulent population is minimal.

**[0018]** Additionally, several nematode-inducible soybean promoters have been identified. Some of these promoters have restricted expression in roots but are highly up-regulated in syncytia. These promoters may be used for targeted RNAi silencing of certain genes. The instant disclosure, together with the newly developed functional analysis tools in soybean such as VIGS (Zhang *et al.*, 2009; Zhang *et al.*, 2010) and the recently completed soybean genome sequencing (Schmutz *et al.*, 2010), may hasten research to understand this relatively unknown, but fascinating, below-ground incompatible plant-pathogen interaction and may ultimately lead to the development of novel strategies to enhance nematode resistance in crop plants.

**[0019]** In one embodiment, one or more of the SCNRGs or fragments thereof may be introduced into a host plant where they are expressed at a level that is higher than the normal expression levels of the same gene(s) in the host plant. The transgenic plant thus generated may be more resistant to soybean cyst nematode (SCN) infection when compared to the host plant. The SCNRG or fragment thereof may encode a protein that is capable of rendering the plant more resistant to SCN infection via a number of different mechanisms. The one or more SCNRGs may be endogenous to the host plant, or they many be exogenous to the host plant.

**[0020]** In another embodiment, the promoters regulating the up-regulation or down-regulation of these SCNRG may be used to control the expression of certain genes. For instance, chimeric construct may be built and introduced into a host plant where the promoter modulates the expression of certain proteins that help render the host plant resistant to SCN. Such construct may contain genes known to play a role in plant defense against SCN infection, or it may contain genes that play a role in an unknown pathway that contribute to SCN resistance. Some of these promoters may be constitutive, others may only be turned on upon detection of SCN invasion by the plant. In another aspect, the promoters may be tissue specific. For example, some promoters may only modulate gene expression in the root tissues. Other promoters may drive a more universal expression of genes in a number of different tissues. These promoters may be used to direct expression of a heterologous gene in a host plant where the heterologous gene encodes a protein that help fight or prevent SCN infection.

**[0021]** In another embodiment, a plant may be modified such that the expression levels of certain SCN responsive genes are altered in a way that render the plant more resistant to SCN infection. For instance, a breeding program may be implemented to select for lines that have elevated levels of one or more of such responsive genes. In one aspect, a method may be

used for generating a transgenic plant that is more resistant to SCN infection from a host plant. This method may include a step of altering the expression levels of a protein encoded by an SCNRG or a fragment thereof, wherein the SCNRG is endogenous to the host plant.

**[0022]** In another embodiment, the expression level of the protein encoded by the SCNRG may be altered so that the level is higher in the transgenic plant than the expression level of the protein in the host plant. In one aspect, the level of the protein encoded by the SCNRG is at least two fold, three fold, or even five fold higher in the transgenic plant than the expression level of the protein in the host plant.

**[0023]** In another embodiment, the expression of the SCNRG may be placed under control of a nematode inducible promoter, such that expression of the SCNRG is induced when the plant is in contact with nematode.

**[0024]** It is to be recognized that not all SCN responsive contribute positively to SCN resistance. As disclosed here, in a SCN resistant line, certain genes may be down-regulated in response to SCN infection. Down-regulation of these genes may contribute to defense against SCN by the resistant lines. In one embodiment, the expression level of certain proteins encoded by certain SCNRGs may be lower in the transgenic plant than the expression level of the protein in the host plant.

**[0025]** In another embodiment, the host plant is a soybean plant that is susceptible to soybean cyst nematode (SCN) infection.

**[0026]** In another embodiment, the expression levels of two more proteins encoded by two or more SCNRGs may be altered in the transgenic plant in order to obtain a transgenic plant that is less susceptible to nematode infection that the host plant. In one aspect, the SCNRG may be any of the disclosed genes that are either up- or down-regulated in the SCN resistant line as compared to the SCN sensitive line. In another aspect, the SCNRG may be GmBAG6, GmAP2, GmBAG6 homolog, GmAP2 homolog, or combination thereof.

## BRIEF DESCRIPTION OF THE DRAWINGS

**[0027]** FIG. 1 shows nematode development and syncytia formation on near-isogenic lines (NIL) of soybean.

**[0028]** FIG. 2 shows the functional classification of differentially expressed genes identified by microarray analysis.

**[0029]** FIG. 3 shows qPCR analysis of up-regulated genes in excised infected whole root pieces of resistant (NIL-R) and susceptible (NIL-S) near-isogenic lines (NIL) at different

days post inoculation (dpi) with avirulent (PA3) or virulent (TN19) soybean cyst nematodes (SCN).

**[0030]** Fig. 4 shows description of promoter-GUS reporter constructs used for microarray validation. (a) Description of probe sets with soybean gene models used for promoter isolation and their putative function; (b) Schematic showing the lengths of promoter elements cloned and their coordinates with respect to the soybean gene model.

**[0031]** Fig. 5 shows the result of promoter-GUS expression in transgenic soybean hairy root lines of resistant near-isogenic lines (NIL-R) infected with PA3 soybean cyst nematodes (SCN).

**[0032]** Fig. 6 shows the result of promoter-GUS expression in transgenic soybean hairy root lines of resistant near-isogenic lines (NIL-R) infected with PA3 soybean cyst nematodes (SCN) using Glyma03g35930.1 (88182p) as the promoter.

**[0033]** Fig. 7 shows the result of promoter-GUS expression in transgenic soybean hairy root lines of resistant near-isogenic lines (NIL-R) infected with PA3 soybean cyst nematodes (SCN) using Glyma03g35920.1(11004p) as the promoter.

**[0034]** Fig. 8 shows longitudinal cross sections of promoter-GUS stained transgenic soybean hairy root lines in the resistant (NIL-R) background infected with PA3 soybean cyst nematodes (SCN).

**[0035]** Fig. 9 shows cell death assay of soybean and Arabidopsis BAG proteins and IQ-BAG domains in yeast *W303-1A* cells.

**[0036]** Fig. 10 shows the range of phenotypes associated with $T_1$ generation of different independent, two week-old *35S* lines overexpressing *GmBag6A* (A-F) and *AtBag6* (G-K) in transgenic Arabidopsis, *Col-0*.

**[0037]** Fig. 11 shows a table summarizing the distribution of phenotypes in Arabidopsis overexpressing BAG6 gene.

**[0038]** Fig. 12 shows that BPMV overexpression of IQ-BAG domain of GmBAG6A (7923R and Glyma07g06750) leads to stunting and cell death phenotype in soybean plants.

**[0039]** Fig. 13 shows that VIGS silencing of GmBAG6 (Glyma07g06750) leads to increased susceptibility of SCN resistant soybean plants to SCN.

**[0040]** Fig. 14 shows that VIGS silencing of GmAP2 transcription factor (Glyma20g029410) leads to increased susceptibility of SCN resistant soybean plants to SCN.

**[0041]** Fig. 15 shows the sequences of the soybean BAG6 genes, GmBAG6A and GmBAG6B, including sequences of cDNA, genomic DNA and the encoded proteins.

## DETAILED DESCRIPTION

[0042] Unless otherwise defined in this disclosure, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the related art. The following terminology and grammatical variants are used in accordance with the definitions set out below.

[0043] The term "genetically altered plant" or "genetically modified plant" refers to a plant whose genetic make-up has been altered or modified such that the modified plant expresses one or more protein that is not normally expressed by the unmodified plant or is expressed at different time or different tissue of the unmodified plant.

[0044] The term "transgenic plant" refers to a host plant into which a gene construct has been introduced. A gene construct, also referred to as a construct, an expression construct, or a DNA construct, generally contains as its components at least a coding sequence and a regulatory sequence. A gene construct typically contains at least one component that is foreign to the host plant. For purpose of this disclosure, all components of a gene construct may be from the host plant, but these components are not arranged in the host in the same manner as they are in the gene construct. A regulatory sequence is a non-coding sequence that typically contribute to the regulation of gene expression, at the transcription or translation levels. It is to be understood that certain segments in the coding sequence may be translated but may be later removed from the functional protein. An example of these segments is the so-called signal peptide, which may facilitate the maturation or localization of the translated protein, but is typically removed once the protein reaches its destination. Examples of a regulatory sequence include but are not limited to a promoter, an enhancer, and certain post-transcriptional regulatory elements.

[0045] After its introduction into a host plant, a gene construct may exist separately from the host chromosomes. Preferably, the entire gene construct, or at least part of it, is integrated onto a host chromosome. The integration may be mediated by a recombination event, which may be homologous, or non-homologous recombination. The term "express" or "expression" refers to production of RNAs using DNAs as template through transcription or translation of proteins from RNAs or the combination of both transcription and translation.

[0046] A "host cell," as used herein, refers to a prokaryotic or eukaryotic cell that contains heterologous DNA which has been introduced into the cell by any means, e.g., electroporation, calcium phosphate precipitation, microinjection, transformation, viral infection, and/or the like. A "host plant" is a plant into which a transgene is to be introduced. A "parental

plant" is the original plant into which genetic changes are to be introduced in order to create a genetically altered plant.

[0047] A "vector" is a composition for facilitating introduction, replication and/or expression of a selected nucleic acid in a cell. Vectors include, for example, plasmids, cosmids, viruses, yeast artificial chromosomes (YACs), etc. A "vector nucleic acid" is a nucleic acid vector into which heterologous nucleic acid is optionally inserted and which can then be introduced into an appropriate host cell. Vectors preferably have one or more origins of replication, and one or more sites into which the recombinant DNA can be inserted. Vectors often have convenient markers by which cells with vectors can be selected from those without. By way of example, a vector may encode a drug resistance gene to facilitate selection of cells that are transformed with the vector. Common vectors include plasmids, phages and other viruses, and "artificial chromosomes." "Expression vectors" are vectors that comprise elements that provide for or facilitate transcription of nucleic acids which are cloned into the vectors. Such elements may include, for example, promoters and/or enhancers operably coupled to a nucleic acid of interest.

[0048] "Plasmids" generally are designated herein by a lower case "p" preceded and/or followed by capital letters and/or numbers, in accordance with standard nomenclatures that are familiar to those of skill in the art. Starting plasmids disclosed herein are either commercially available, publicly available on an unrestricted basis, or can be constructed from available plasmids by routine application of well known, published procedures. Many plasmids and other cloning and expression vectors are well known and readily available to those of skill in the art. Moreover, those of skill readily may construct any number of other plasmids suitable for use as described below. The properties, construction and use of such plasmids, as well as other vectors, is readily apparent to those of ordinary skill upon reading the present disclosure.

[0049] When a molecule is identified in or can be isolated from a organism, it can be said that such a molecule is derived from said organism. When two organisms have significant difference in the genetic materials in their respective genomes, these two organisms can be said to be genetically different. For purpose of this disclosure, the term "plant" means a whole plant, a seed, or any organ or tissue of a plant that may potentially deveolop into a whole plant.

[0050] The term "isolated" means that the material is removed from its original environment, such as the native or natural environment if the material is naturally occurring. For example, a naturally-occurring nucleic acid, polypeptide, or cell present in a living animal is not isolated, but the same polynucleotide, polypeptide, or cell separated from some or all of the coexisting materials in the natural system, is isolated, even if subsequently reintroduced into the

natural system. Such nucleic acids can be part of a vector and/or such nucleic acids or polypeptides could be part of a composition, and still be isolated in that such vector or composition is not part of its natural environment.

**[0051]** A "recombinant nucleic acid" is one that is made by recombining nucleic acids, e.g., during cloning, DNA evolution or other procedures. A "recombinant polypeptide" is a polypeptide which is produced by expression of a recombinant nucleic acid. An "amino acid sequence" is a polymer of amino acid residues (a protein, polypeptide, etc.) or a character string representing an amino acid polymer, depending on context. Either the given nucleic acid or the complementary nucleic acid can be determined from any specified polynucleotide sequence.

**[0052]** The terms "nucleic acid," or "polynucleotide" refer to a deoxyribonucleotide, in the case of DNA ,or ribonucleotide in the case of RNA polymer in either single- or double-stranded form, and unless otherwise specified, encompasses known analogues of natural nucleotides that can be incorporated into nucleic acids in a manner similar to naturally occurring nucleotides. A "polynucleotide sequence" is a nucleic acid which is a polymer of nucleotides (A,C,T,U,G, etc. or naturally occurring or artificial nucleotide analogues) or a character string representing a nucleic acid, depending on context. Either the given nucleic acid or the complementary nucleic acid can be determined from any specified polynucleotide sequence.

**[0053]** A "subsequence" or "fragment" is any portion of an entire sequence of a DNA, RNA or polypeptide molecule, up to and including the complete sequence. Typically a subsequence or fragment comprises less than the full-length sequence, and is sometimes referred to as the "truncated version."

**[0054]** Nucleic acids and/or nucleic acid sequences are "homologous" when they are derived, naturally or artificially, from a common ancestral nucleic acid or nucleic acid sequence. Proteins and/or protein sequences are homologous when their encoding DNAs are derived, naturally or artificially, from a common ancestral nucleic acid or nucleic acid sequence. Similarly, nucleic acids and/or nucleic acid sequences are homologous when they are derived, naturally or artificially, from a common ancestral nucleic acid or nucleic acid sequence. The homologous molecules can be termed homologs. Homology is generally inferred from sequence identity between two or more nucleic acids or proteins (or sequences thereof). The precise percentage of identity between sequences that is useful in establishing homology varies with the nucleic acid and protein at issue, but as little as 25% sequence identity is routinely used to establish homology. Higher levels of sequence identity, e.g., 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95% or 99% or more can also be used to establish homology. Methods for determining

sequence identity percentages (e.g., BLASTP and BLASTN using default parameters) are described herein and are generally available.

**[0055]** The terms "identical" or "sequence identity" in the context of two nucleic acid sequences or amino acid sequences of polypeptides refers to the residues in the two sequences which are the same when aligned for maximum correspondence over a specified comparison window. A "comparison window", as used herein, refers to a segment of at least about 20 contiguous positions, usually about 50 to about 200, more usually about 100 to about 150 in which a sequence may be compared to a reference sequence of the same number of contiguous positions after the two sequences are aligned optimally. Methods of alignment of sequences for comparison are well-known in the art. Optimal alignment of sequences for comparison may be conducted by the local homology algorithm of Smith and Waterman (1981) *Adv. Appl. Math.* 2:482; by the alignment algorithm of Needleman and Wunsch (1970) *J. Mol. Biol.* 48:443; by the search for similarity method of Pearson and Lipman (1988) *Proc. Nat. Acad. Sci U.S.A.* 85:2444; by computerized implementations of these algorithms (including, but not limited to CLUSTAL in the PC/Gene program by Intelligentics, Mountain View Calif., GAP, BESTFIT, BLAST, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group (GCG), 575 Science Dr., Madison, Wis., U.S.A.); the CLUSTAL program is well described by Higgins and Sharp (1988) *Gene* 73:237-244 and Higgins and Sharp (1989) *CABIOS* 5:151-153; Corpet et al. (1988) *Nucleic Acids Res.* 16:10881-10890; Huang et al (1992) *Computer Applications in the Biosciences* 8:155-165; and Pearson et al. (1994) *Methods in Molecular Biology* 24:307-331. Alignment is also often performed by inspection and manual alignment.

**[0056]** In one class of embodiments, the polypeptides herein are at least 70%, generally at least 75%, optionally at least 80%, 85%, 90%, 98% or 99% or more identical to a reference polypeptide, e.g., those that are encoded by DNA sequences as set forth by any one of the SCNRGs disclosed herein or a fragment thereof, e.g., as measured by BLASTP (or CLUSTAL, or any other available alignment software) using default parameters. Similarly, nucleic acids can also be described with reference to a starting nucleic acid, e.g., they can be 50%, 60%, 70%, 75%, 80%, 85%, 90%, 98%, 99% or more identical to a reference nucleic acid, e.g., those that are set forth by any one of the SCNRGs disclosed herein or a fragment thereof, e.g., as measured by BLASTN (or CLUSTAL, or any other available alignment software) using default parameters. When one molecule is said to have certain percentage of sequence identity with a larger molecule, it means that when the two molecules are optimally aligned, said

percentage of residues in the smaller molecule finds a match residue in the larger molecule in accordance with the order by which the two molecules are optimally aligned.

**[0057]** The term "substantially identical" as applied to nucleic acid or amino acid sequences means that a nucleic acid or amino acid sequence comprises a sequence that has at least 90% sequence identity or more, preferably at least 95%, more preferably at least 98% and most preferably at least 99%, compared to a reference sequence using the programs described above (preferably BLAST) using standard parameters. For example, the BLASTN program (for nucleotide sequences) uses as defaults a word length (W) of 11, an expectation (E) of 10, M=5, N=-4, and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a word length (W) of 3, an expectation (E) of 10, and the BLOSUM62 scoring matrix (*see* Henikoff & Henikoff, *Proc. Natl. Acad. Sci. USA* 89:10915 (1989)). Percentage of sequence identity is determined by comparing two optimally aligned sequences over a comparison window, wherein the portion of the polynucleotide sequence in the comparison window may comprise additions or deletions (i.e., gaps) as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical nucleic acid base or amino acid residue occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison and multiplying the result by 100 to yield the percentage of sequence identity. Preferably, the substantial identity exists over a region of the sequences that is at least about 50 residues in length, more preferably over a region of at least about 100 residues, and most preferably the sequences are substantially identical over at least about 150 residues. In a most preferred embodiment, the sequences are substantially identical over the entire length of the coding regions.

**[0058]** The term "polypeptide" is used interchangeably with the terms "polypeptides" and "protein(s)", and refers to a polymer of amino acid residues. A 'mature protein' is a protein which is full-length and which, optionally, includes glycosylation or other modifications typical for the protein in a given cellular environment.

**[0059]** The term "variant" or "mutant" with respect to a polypeptide refers to an amino acid sequence that is altered by one or more amino acids with respect to a reference sequence. The variant may have "conservative" changes, wherein a substituted amino acid has similar structural or chemical properties, e.g., replacement of leucine with isoleucine. Alternatively, a variant may have "nonconservative" changes, e.g., replacement of a glycine with a tryptophan. Analogous minor variation can also include amino acid deletion or insertion,

or both. Guidance in determining which amino acid residues can be substituted, inserted, or deleted without eliminating biological or immunological activity can be found using computer programs well known in the art, for example, DNASTAR software.

**[0060]** A variety of additional terms are defined or otherwise characterized herein. In practicing the instrumentalities described herein, many conventional techniques in molecular biology, microbiology, and recombinant DNA are optionally used. These techniques are well known to those of ordinary skill in the art. For example, one skilled in the art would be familiar with techniques for *in vitro* amplification methods, including the polymerase chain reaction (PCR), for the production of the homologous nucleic acids described herein.

**[0061]** In addition, commercially available kits may facilitate the purification of plasmids or other relevant nucleic acids from cells. See, for example, EasyPrep™ and FlexiPrep™ kits, both from Pharmacia Biotech; StrataClean™ from Stratagene; and, QIAprep™ from Qiagen. Any isolated and/or purified nucleic acid can be further manipulated to produce other nucleic acids, used to transfect cells, incorporated into related vectors to infect organisms, or the like. Typical cloning vectors contain transcription terminators, transcription initiation sequences, and promoters useful for regulation of the expression of the particular target nucleic acid. The vectors optionally comprise generic expression cassettes containing at least one independent terminator sequence, sequences permitting replication of the cassette in eukaryotes, or prokaryotes, or both, (e.g., shuttle vectors) and selection markers for both prokaryotic and eukaryotic systems. Vectors are suitable for replication and integration in prokaryotes, eukaryotes, or both.

**[0062]** Various types of mutagenesis are optionally used to modify a gene and their encoded polypeptides, as described herein, to produce conservative or non-conservative variants. Any available mutagenesis procedure can be used. Such mutagenesis procedures optionally include selection of mutant nucleic acids and polypeptides for one or more activity of interest. Procedures that can be used include, but are not limited to: site-directed point mutagenesis, random point mutagenesis, *in vitro* or *in vivo* homologous recombination (DNA shuffling), mutagenesis using uracil-containing templates, oligonucleotide-directed mutagenesis, phosphorothioate-modified DNA mutagenesis, mutagenesis using gapped duplex DNA, point mismatch repair, mutagenesis using repair-deficient host strains, restriction-selection and restriction-purification, deletion mutagenesis, mutagenesis by total gene synthesis, double-strand break repair, mutagenesis by chimeric constructs, and many others known to persons of skill in the art.

**[0063]** In one embodiment, mutagenesis can be guided by known information about the naturally occurring molecule or altered or mutated naturally occurring molecule. By way of example, this known information may include sequence, sequence comparisons, physical properties, crystal structure and the like. In another class of mutagenesis, modification is essentially random, e.g., as in classical DNA shuffling.

**[0064]** Polypeptides may include variants, in which the amino acid sequence has at least 70% identity, preferably at least 80% identity, typically 90% identity, preferably at least 95% identity, more preferably at least 98% identity and most preferably at least 99% identity, to the amino acid sequences as encoded by the DNA sequences set forth in any one of the SCN responsive genes disclosed herein.

**[0065]** The aforementioned polypeptides may be obtained by any of a variety of methods. Smaller peptides (less than 50 amino acids long) are conveniently synthesized by standard chemical techniques and can be chemically or enzymatically ligated to form larger polypeptides. Polypeptides can be purified from biological sources by methods well known in the art, for example, as described in *Protein Purification, Principles and Practice, Second Edition* Scopes, Springer Verlag, N.Y. (1987) Polypeptides are optionally but preferably produced in their naturally occurring, truncated, or fusion protein forms by recombinant DNA technology using techniques well known in the art. These methods include, for example, *in vitro* recombinant DNA techniques, synthetic techniques and *in vivo* genetic recombination. See, for example, the techniques described in Sambrook et al. (2001) *Molecular Cloning, A Laboratory Manual*, Third Edition, Cold Spring Harbor Press, N.Y.; and Ausubel et al., eds. (1997) *Current Protocols in Molecular Biology*, Green Publishing Associates, Inc., and John Wiley & Sons, Inc., N.Y (supplemented through 2002). RNA encoding the proteins may also be chemically synthesized. See, for example, the techniques described in *Oligonucleotide Synthesis*, (1984) Gait ed., IRL Press, Oxford, which is incorporated by reference herein in its entirety.

**[0066]** The nucleic acid molecules described herein may be expressed in a suitable host cell or an organism to produce proteins. Expression may be achieved by placing a nucleotide sequence encoding these proteins into an appropriate expression vector and introducing the expression vector into a suitable host cell, culturing the transformed host cell under conditions suitable for expression of the proteins described or variants thereof, or a polypeptide that comprises one or more domains of such proteins. The recombinant proteins from the host cell may be purified to obtain purified and, preferably, active protein.

Alternatively, the expressed protein may be allowed to function in the intact host cell or host organism.

**[0067]** Appropriate expression vectors are known in the art, and may be purchased or applied for use according to the manufacturer's instructions to incorporate suitable genetic modifications. For example, pET-14b, pcDNA1Amp, and pVL1392 are available from Novagen and Invitrogen, and are suitable vectors for expression in *E. coli*, mammalian cells and insect cells, respectively. These vectors are illustrative of those that are known in the art, and many other vectors can be used for the same purposes. Suitable host cells can be any cell capable of growth in a suitable media and allowing purification of the expressed protein. Examples of suitable host cells include bacterial cells, such as *E. coli*, *Streptococci*, *Staphylococci*, *Streptomyces* and *Bacillus subtilis* cells; fungal cells such as *Saccharomyces* and *Aspergillus* cells; insect cells such as *Drosophila* S2 and *Spodoptera* Sf9 cells, mammalian cells such as CHO, COS, HeLa, 293 cells; and plant cells.

**[0068]** Culturing and growth of the transformed host cells can occur under conditions that are known in the art. The conditions will generally depend upon the host cell and the type of vector used. Suitable culturing conditions may be used such as temperature and chemicals and will depend on the type of promoter utilized.

**[0069]** Purification of the proteins or domains of such proteins, if desired, may be accomplished using known techniques without performing undue experimentation. Generally, the transformed cells expressing one of these proteins are broken, crude purification occurs to remove debris and some contaminating proteins, followed by chromatography to further purify the protein to the desired level of purity. Host cells may be broken by known techniques such as homogenization, sonication, detergent lysis and freeze-thaw techniques. Crude purification can occur using ammonium sulfate precipitation, centrifugation or other known techniques. Suitable chromatography includes anion exchange, cation exchange, high performance liquid chromatography (HPLC), gel filtration, affinity chromatography, hydrophobic interaction chromatography, etc. Well known techniques for refolding proteins can be used to obtain the active conformation of the protein when the protein is denatured during intracellular synthesis, isolation or purification.

**[0070]** Sequence information of the SCN responsive genes may also be used to design oligonucleotides for detecting their mRNA levels in the cells or in plant tissues. For example, the oligonucleotides can be used in a Northern blot analysis to quantify the levels of the mRNA. Moreover, full-length or fragment of the SCN responsive genes may be used in preparing microarrays (or gene chips). Full-length or fragment of the SCN responsive genes

may also be used in microarray experiments to study expression profile of the SCN responsive genes. High-throughput screening can be conducted to measure expression levels of the SCN responsive genes in different cells or tissues. Various compounds or other external factors may be screened for their effects expression of the SCN responsive gene expression.

[0071] Sequences of the SCN responsive genes and proteins identified herein may also provide a tool for identification of other proteins that may be involved in plant defense against SCN. For example, chimeric SCN resistant proteins can be used as a "bait" to identify other proteins that interact with SCN resistant proteins in a yeast two-hybrid screening. Recombinant SCN resistant proteins can also be used in pull-down experiment to identify their interacting proteins. These other proteins may be cofactors that enhance the function of the SCN resistant proteins, or they may be SCN resistant proteins themselves which have not been identified in the experiments disclosed herein.

[0072] The SCN resistant polypeptides may possess structural features which can be recognized, for example, by using immunological assays. The generation of antisera which specifically bind the SCN resistant polypeptides, as well as the polypeptides which are bound by such antisera, are a feature of the disclosed embodiments.

[0073] In order to produce antisera for use in an immunoassay, one or more of the immunogenic SCN resistant polypeptides or fragments thereof are produced and purified as described herein. For example, recombinant protein may be produced in a host cell such as a bacterial or an insect cell. The resultant proteins can be used to immunize a host organism in combination with a standard adjuvant, such as Freund's adjuvant. Commonly used host organisms include rabbits, mice, rats, donkeys, chickens, goats, horses, etc. An inbred strain of mice may also be used to obtain more reproducible results due to the virtual genetic identity of the mice. The mice are immunized with the immunogenic SCNRG polypeptides in combination with a standard adjuvant, such as Freund's adjuvant, and a standard mouse immunization protocol. See, for example, Harlow and Lane, *Antibodies, A Laboratory Manual*, Cold Spring Harbor Publications, New York (1988), which provides comprehensive descriptions of antibody generation, immunoassay formats and conditions that can be used to determine specific immunoreactivity. Alternatively, one or more synthetic or recombinant SCN resistant polypeptides or fragments thereof derived from the sequences disclosed herein is conjugated to a carrier protein and used as an immunogen.

[0074] Antisera that specifically bind the SCN resistant proteins may be used in a range of applications, including but not limited to immunofluorescence staining of cells for the

expression level and localization of the SCN resistant proteins, cytological staining for the expression of SCN resistant proteins in tissues, as well as in Western blot analysis.

[0075] Another aspect of the disclosure includes screening for potential or candidate modulators of SCN resistant protein activity. For example, potential modulators may include small molecules, organic molecules, inorganic molecules, proteins, hormones, transcription factors, or the like, which can be contacted to a cell or certain tissues that express the SCN resistant proteins to assess the effects, if any, of the candidate modulator upon SCN resistant protein activity.

[0076] Alternatively, candidate modulators may be screened to modulate expression of SCN resistant proteins. For example, potential modulators may include small molecules, organic molecules, inorganic molecules, proteins, hormones, transcription factors, or the like, which can be contacted to a cell or certain tissues that express the SCN resistant proteins, to assess the effects, if any, of the candidate modulator upon SCN resistant protein expression. Expression of a SCN responsive gene described herein may be detected, for example, via Northern blot analysis or quantitative (optionally real time) RT-PCR, before and after application of potential expression modulators. Alternatively, promoter regions of the various SCN responsive genes may be coupled to reporter constructs including, without limitation, CAT, beta-galactosidase, luciferase or any other available reporter, and may similarly be tested for expression activity modulation by the candidate modulator. Promoter regions of the various genes are generally sequences in the proximity upstream of the start site of transcription, typically within 1 Kb or less of the start site, such as within 500 bp, 250 bp or 100 bp of the start site. In certain cases, a promoter region may be located between 1 and 5 Kb from the start site.

[0077] In either case, whether the assay is to detect modulated activity or expression, a plurality of assays may be performed in a high-throughput fashion, for example, using automated fluid handling and/or detection systems in serial or parallel fashion. Similarly, candidate modulators can be tested by contacting a potential modulator to an appropriate cell using any of the activity detection methods herein, regardless of whether the activity that is detected is the result of activity modulation, expression modulation or both.

[0078] A method of modifying a plant may include introducing into a host plant one or more SCN responsive genes described above. The SCN responsive genes may be placed in an expression construct, which may be designed such that the SCN resistant protein(s) are expressed constitutively, or inducibly. The construct may also be designed such that the SCN resistant protein(s) are expressed in certain tissue(s), but not in other tissue(s). The SCN resistant protein(s) may enhance the ability of the host plant to defend SCN infection. The host

plant may include any plants whose growth and/or yield may be enhanced by a modified SCN response. Methods for generating such transgenic plants is well known in the field. *See e.g.*, Leandro Peña (Editor), Transgenic Plants: Methods and Protocols (Methods in Molecular Biology), Humana Press, 2004.

**[0079]** The use of gene inhibition technologies such as antisense RNA, antificial microRNA, or co-suppression or double stranded RNA interference is also within the scope of the present disclosure. In these approaches, the isolated gene sequence is operably linked to a suitable regulatory element. In one embodiment of the disclosure, the construct contains a DNA expression cassette that contains, in addition to the DNA sequences required for transformation and selection in said cells, a DNA sequence that encodes a SCN resistant proteins or a SCN resistant modulator protein, with at least a portion of said DNA sequence in an antisense orientation relative to the normal presentation to the transcriptional regulatory region, operably linked to a suitable transcriptional regulatory region such that said recombinant DNA construct expresses an antisense RNA or portion thereof of an antisense RNA in the resultant transgenic plant.

**[0080]** It is apparent to one of skill in the art that the polynucleotide encoding the SCN resistant proteins or a SCN resistant modulator proteins can be in the antisense (for inhibition by antisense RNA) or sense (for inhibition by co-suppression) orientation, relative to the transcriptional regulatory region. Alternatively a combination of sense and antisense RNA expression can be utilized to induce double stranded RNA interference. *See, e.g.*, Chuang and Meyerowitz, PNAS 97: 4985-4990, 2000; *see also* Smith et al., Nature 407: 319-320, 2000.

**[0081]** These methods for generation of transgenic plants generally entail the use of transformation techniques to introduce the gene or construct encoding the SCN resistant proteins or a SCN resistant modulator proteins, or a part or a homolog thereof, into plant cells. Transformation of a plant cell can be accomplished by a variety of different methodology. Methods that have general utility include, for example, Agrobacterium based systems, using either binary and/or cointegrate plasmids of both A. tumifaciens and A. rhyzogenies, (*See e.g.*, U.S. Pat. No. 4,940,838, U.S. Pat. No. 5,464,763), the biolistic approach (See e.g, U.S. Pat. No. 4,945,050, U.S. Pat. No. 5,015,580, U.S. Pat. No. 5,149,655), microinjection, (See e.g., U.S. Pat. No. 4,743,548), direct DNA uptake by protoplasts, (See e.g., U.S. Pat. No. 5,231,019, U.S. Pat. No. 5,453,367) or needle-like whiskers (See e.g., U.S. Pat. No. 5,302,523). Any method for the introduction of foreign DNA into a plant cell and for expression therein may be used within the context of the present disclosure.

[0082] Plants that are capable of being transformed encompass a wide range of species, including but not limited to soybean, corn, potato, rice, wheat and many other crops, fruit plants, vegetables and tobacco. *See generally*, Vain, P., Thirty years of plant transformation technology development, Plant Biotechnol J. 2007 Mar;5(2):221-9. Any plants that are capable of taking in foreign DNA and transcribing the DNA into RNA and/or further translating the RNA into a protein may be a suitable host.

[0083] The modulators described above that may alter the expression levels or the activity of the SCN resistant proteins (collectively called SCN resistant modulators) may also be introduced into a host plant in the same or similar manner as described above. In one embodiment, the SCN resistant modulators are primarily transcription factors that regulate the transcription of the SCN responsive genes.

[0084] The SCN resistant proteins or the SCN resistant modulators may be used to modify a target plant by causing them to be assimilated by the plant. Alternatively, the SCN resistant proteins or the SCN resistant modulators may be applied to a target plant by causing them to be in contact with the plant, or with a specific organ or tissue of the plant. In one embodiment, organic or inorganic molecules that can function as SCN resistant modulators may be caused to be in contact with a plant such that these chemicals may enhance defense against SCN by the target plant.

[0085] In addition to the SCN resistant modulators, SCN resistant polypeptides or SCN resistant nucleic acids, a composition containing other ingredients may also be introduced, administered or delivered to the plant to be modified. In one aspect, a composition containing an agriculturally acceptable ingredient may be used in conjunction with the SCN resistant modulators to be administered or delivered to the plant.

[0086] Bioinformatic systems are widely used in the art, and can be utilized to identify homology or similarity between different character strings, or can be used to perform other desirable functions such as to control output files, provide the basis for making presentations of information including the sequences and the like. Examples include BLAST, discussed *supra*. For example, commercially available databases, computers, computer readable media and systems may contain character strings corresponding to the sequence information herein for the SCN resistant polypeptides and nucleic acids described herein. These sequences may include specifically the SCN resistant sequences listed herein and the various silent substitutions and conservative substitutions thereof.

[0087] The bioinformatic systems contain a wide variety of information that includes, for example, a complete sequence listings for the entire genome of an individual organism

representing a species. Thus, for example, using the SCN resistant sequences as a basis for comparison, the bioinformatic systems may be used to compare different types of homology and similarity of various stringency and length on the basis of reported data. These comparisons are useful to identify homologs or orthologs where, for example, the basic SCNRG gene ortholog is shown to be conserved across different organisms. Thus, the bioinformatic systems may be used to detect or recognize the homologs or orthologs, and to predict the function of recognized homologs or orthologs. By way of example, many homology determination methods have been designed for comparative analysis of sequences of biopolymers including nucleic acids, proteins, etc.. With an understanding of hydrogen bonding between the principal bases in natural polynucleotides, models that simulate annealing of complementary homologous polynucleotide strings can also be used as a foundation of sequence alignment or other operations typically performed on the character strings corresponding to the sequences herein. One example of a software package for calculating sequence similarity is BLAST, which can be adapted to the present invention by inputting character strings corresponding to the sequences herein.

**[0088]** The software can also include output elements for controlling nucleic acid synthesis (*e.g.*, based upon a sequence or an alignment of a sequences herein) or other operations which occur downstream from an alignment or other operation performed using a character string corresponding to a sequence herein.

**EXAMPLES**

**[0089]** The following nonlimiting examples report general procedures, reagents and characterization methods that teach by way of example, and should not be construed in a narrowing manner that limits the disclosure to what is specifically disclosed. Those skilled in the art will understand that numerous modifications may be made and still the result will fall within the spirit and scope of the present invention.

**Plant and nematode material**

**[0090]** Seeds of soybean (*Glycine max* (L.) Merr) near-isogenic lines (NIL) differing at the *Rhg1* locus (NIL-R and NIL-S) were derived from a cross between the resistance source PI 209332 and the susceptible cultivar Evans (Mudge, 1999). The SCN (*Heterodera glycines*, Ichinohe) inbred populations PA3 and TN19 were obtained from a publicly available collection at the University of Illinois at Urbana-Champaign and mass-selected according to standard procedures (Niblack *et al.*, 1993) on soybean cv. Williams 82 and PI 437654, respectively. HG-

type tests (Niblack *et al.*, 2002) confirmed that the PA3 population was HG-type 0 and the TN19 population was HG-type 1-7.

**Laser capture microdissection**

**[0091]** PA3 and TN19-infected root pieces (~ 1cm) of the NIL-R and NIL-S were excised at 5 dpi or 8 dpi and immediately processed for laser capture microdissection according to Ithal *et al.* (2007b).

**Microarray hybridization, statistical analysis, and qPCR validation**

**[0092]** RNA extraction, amplification, and labeling were performed according to Ithal *et al.* (2007b). The samples were sent to the Iowa State University GeneChip microarray core facility for fragmentation, hybridization, staining, and scanning of the GeneChip Soybean Genome Array (Affymetrix). The logarithms of the Affymetrix MAS 5.0 signals were normalized by computing the median of the log signals on each chip and then aligning these medians to a common value. These normalized expression data were analyzed on a gene-by-gene basis using SAS. Each analysis was based on a randomized complete block design with three replications as blocks and the four combinations of genotype (resistant versus susceptible) and days post infection (5 dpi and 8 dpi) as treatments. Tests for genotype main effects, dpi main effects, and genotype by dpi interaction were conducted for each gene. The resulting p-values were converted to q-values as described by Storey and Tibshirani (2003). These q-values were used to control the estimated False Discovery Rate (FDR) at desired levels. For example, by declaring differential expression between resistant and susceptible genotypes for all genes with q-values less than or equal to 0.10, the proportion of false positives among all genes declared differentially expressed is expected to be approximately 10%. Annotations and classifications were based on SoyBase Affymetrix™ GeneChip® Soybean Genome Array Annotation, Version 2 (http://soybase.org/AffyChip/). Arabidopsis unique gene identifiers (At-numbers) were downloaded from The Arabidopsis Information Resource (TAIR, www.arabidopsis.org) for the top hits. The microarray data are deposited in the ArrayExpress database at the European Bioinformatics Institute under accession number X (submitted). qPCR validation studies were conducted according to Ithal *et al* (2007b).

**[0093]** All sequence in Tables 1-17 contain soybean genome database sequences that are publicly available and are identified by a unique sequence identifier in those databases, which are hereby incorporated by reference into this disclosure.

**Table 1 Candidate genes related to the structure or function of the cell wall**

| Affy ID | Q-value | Fold change | Soybean Gene | Arabidopsis homolog | Description |
|---|---|---|---|---|---|
| GmaAffx.33157.1.S1_at | 0.0317164 | 9.28486499 | Glyma14g07550.1 | AT4G13340.1 | leucine-rich repeat family protein / extensin family protein FUNCTIONS IN: structural constituent of cell wall |
| Gma.8565.1.S1_at | 0.0272696 | 8.316879514 | Glyma13g00280.1 | AT4G25810.1 | xyloglucan endotransglycosylase-related protein (XTR6) |
| Gma.8077.1.S1_at | 0.0163886 | 4.589862031 | No Soybean Match Identified | AT3G23730.1 | xyloglucan: xyloglucosyl transferase, putative / xyloglucan endotransglycosylase, putative / endo-xyloglucan transferase, putative; FUNCTIONS IN: hydrolase activity, acting on glycosyl bonds, hydrolase activity, hydrolyzing O-glycosyl compounds, xyloglucan:xyloglucosyl transferase activity; INVOLVED IN: carbohydrate metabolic process, cellular glucan metabolic process; LOCATED IN: endomembrane system, cell wall, apoplast; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Xyloglucan endotransglucosylase/hydrolase (InterPro:IPR016455), Xyloglucan endo-transglycosylase, C-terminal (InterPro:IPR010713), Concanavalin A-like lectin/glucanase, subgroup (InterPro:IPR013320), Concanavalin A-like lectin/glucanase (InterPro:IPR008985), Glycoside hydrolase, family 16, active site (InterPro:IPR008263), Glycoside hydrolase, family 16 (InterPro:IPR000757); BEST Arabidopsis thaliana protein match is: XTR7 (XYLOGLUCAN ENDOTRANSGLYCOSYLASE 7); |
| GmaAffx.80271.1.S1_at | 0.0085675 | 3.76828919 | Glyma13g35100.1 | AT2G27380.1 | Encodes an extensin like gene involved in seed germination. |

| | | | | | |
|---|---|---|---|---|---|
| Gma.2763.2.S1_at | 0.0998527 | 3.653971559 | Glyma18g09490.1 | AT5G55500.1 | Encodes a beta-1,2-xylosyltransferase that is glycosylated at two positions. |
| Gma.117.1.S1_at | 0.0163687 | 3.280787916 | Glyma17g07260.1 | AT4G25810.1 | xyloglucan endotransglycosylase-related protein (XTR6) |
| Gma.4568.1.A1_at | 0.0963033 | 3.04499165 | Glyma17g03390.1 | AT4G16120.1 | putative membrane-anchored cell wall protein |
| GmaAffx.11502.1.S1_at | 0.0362282 | 2.130359066 | Glyma16g01640.1 | AT1G53840.1 | encodes a pectin methylesterase |
| Gma.5312.2.S1_a_at | 0.0771246 | 1.898262486 | Glyma11g10240.4 | AT3G45970.1 | member of EXPANSIN-LIKE. |
| Gma.12505.2.A1_at | 0.0968339 | 1.881230191 | Glyma09g05010.2 | AT4G02500.1 | Encodes a protein with xylosyltransferase activity, which is specific for UDP-xylose as donor substrate and for oligosaccharides with a degree of polymerization >4. Although the enzyme utilizes either cellopentaose or cellohexaose, its activity is four-fold higher with cellohexaose as an acceptor compared to cellopentaose. The enzyme is able to add several xylosyl residues to the acceptor forming mono-, di- and trixylosylated polysaccharides. |
| GmaAffx.82513.1.S1_at | 0.0436431 | -4.18849273 | Glyma07g07280.1 | AT3G61490.2 | glycoside hydrolase family 28 protein / polygalacturonase (pectinase) family protein; FUNCTIONS IN: polygalacturonase activity |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.37897.1.A1_at | 0.0565355 | -4.00276357 | Glyma16g03680.1 | AT3G61490.2 | glycoside hydrolase family 28 protein / polygalacturonase (pectinase) family protein; FUNCTIONS IN: polygalacturonase activity; INVOLVED IN: carbohydrate metabolic process; LOCATED IN: endomembrane system; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Pectin lyase fold/virulence factor (InterPro:IPR011050), Glycoside hydrolase, family 28 (InterPro:IPR000743), Pectin lyase fold (InterPro:IPR012334), Parallel beta-helix repeat (InterPro:IPR006626); BEST Arabidopsis thaliana protein match is: glycoside hydrolase family 28 protein / polygalacturonase (pectinase) family protein (TAIR:AT4G23500.1); Has 2372 Blast hits to 2368 proteins in 307 species: Archae - 2; Bacteria - 553; Metazoa - 8; Fungi - 879; Plants - 827; Viruses - 0; Other Eukaryotes - 103 (source: |
| Gma.4216.2.S1_s_at | 0.0581114 | -2.94662047 | Glyma16g26630.1 | AT5G13870.1 | EXGT-A4, endoxyloglucan transferase, |
| Gma.12828.1.A1_at | 0.0581114 | -2.49416625 | Glyma05g04190.1 | AT4G12390.1 | PECTIN METHYLESTERASE INHIBITOR 1 (PME1) |
| GmaAffx.67477.2.S1_at | 0.0673275 | -2.42933853 | Glyma09g37090.2 | AT1G69530.2 | Member of Alpha-Expansin Gene Family. Naming convention from the Expansin Working Group (Kende et al, Plant Mol Bio). Involved in the formation of nematode-induced syncytia in roots of Arabidopsis thaliana. |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.34407.1.S1_at | 0.0669662 | -2.32524781 | Glyma17g35760.1 | AT4G06744.1 | leucine-rich repeat family protein / extensin family protein; FUNCTIONS IN: protein binding; INVOLVED IN: biological_process unknown; CONTAINS InterPro DOMAIN/s: Leucine-rich repeat (InterPro:IPR001611); BEST Arabidopsis thaliana protein match is: leucine-rich repeat family protein (TAIR:AT3G19320.1); Has 25956 Blast hits to 7632 proteins in 409 species: Archae - 10; Bacteria - 542; Metazoa - 1041; Fungi - 73; Plants - 22939; Viruses - 0; Other Eukaryotes - 1351 (source: NCBI BLink). |
| Gma.16913.1.S1_s_at | 0.0631501 | -2.20500841 | Glyma09g12200.7 | AT2G27380.1 | Encodes an extensin like gene involved in seed germination. |
| GmaAffx.17518.1.S1_at | 0.07755 | -2.17208917 | Glyma03g38350.3 | AT3G62110.1 | glycoside hydrolase family 28 protein / polygalacturonase (pectinase) family protein; FUNCTIONS IN: polygalacturonase activity |
| Gma.576.1.S1_at | 0.0957992 | -2.14205359 | Glyma06g47740.1 | AT5G62350.1 | invertase/pectin methylesterase inhibitor family protein / DC 1.2 homolog (FL5-2I22); pectinesterase inhibitor activity |
| GmaAffx.18757.1.S1_at | 0.0596157 | -2.12959892 | Glyma09g39930.1 | AT5G62150.1 | peptidoglycan-binding LysM domain-containing protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: cell wall macromolecule catabolic process; LOCATED IN: endomembrane system; EXPRESSED IN: 15 plant structures; EXPRESSED DURING: 6 growth stages; CONTAINS InterPro DOMAIN/s: Peptidoglycan-binding lysin domain (InterPro:IPR018392); BEST Arabidopsis thaliana protein match is: peptidoglycan-binding LysM domain-containing protein (TAIR:AT4G25433.1); Has 30 Blast hits to 30 proteins in 7 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 30; Viruses - 0; Other Eukaryotes - 0 |

| | | | | | |
|---|---|---|---|---|---|
| Gma.15996.1.S1_at | 0.0751423 | -1.64105397 | Glyma15g23830.2 | AT2G27380.1 | Encodes an extensin like gene involved in seed germination. |
| GmaAffx.78421.1.S1_at | 0.0498357 | -1.57602812 | Glyma08g15650.1 | AT5G09760.1 | pectinesterase family protein; FUNCTIONS IN: enzyme inhibitor activity, pectinesterase activity; INVOLVED IN: cell wall modification; LOCATED IN: cell wall, chloroplast, plant-type cell wall; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Pectin lyase fold/virulence factor (InterPro:IPR011050), Pectinesterase inhibitor (InterPro:IPR006501), Pectinesterase, catalytic (InterPro:IPR000070), Pectin lyase fold (InterPro:IPR012334); BEST Arabidopsis thaliana protein match is: pectinesterase family protein (TAIR:AT5G64640.1); Has 1447 Blast hits to 1415 proteins in 271 species: Archae - 0; Bacteria - 416; Metazoa - 1; Fungi - 117; Plants - |
| GmaAffx.6824.2.S1_at | 0.0973423 | -1.52866562 | Glyma08g04860.1 | AT1G14890.1 | enzyme inhibitor/ pectinesterase/ pectinesterase inhibitor |
| Gma.79.4.S1_s_at | 0.0851155 | -1.49752901 | Glyma09g12200.7 | AT2G27380.1 | Encodes an extensin like gene involved in seed germination. |

**Table 2 Candidate genes related to protein sorting and transport**

| Affy ID | Q-value | Fold change | Soybean Gene | Arabidopsis homolog | Description |
|---|---|---|---|---|---|
| GmaAffx.35140.1.S1_at | 0.0267886 | 45.9674 | Glyma04g11370.1 | AT1G22180.2 | SEC14 cytosolic factor family protein / phosphoglyceride transfer family protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 19 plant structures; EXPRESSED DURING: 9 growth stages; CONTAINS InterPro DOMAIN/s: Cellular retinaldehyde-binding/triple function, C-terminal (InterPro:IPR001251), Cellular retinaldehyde-binding/triple function, N-terminal (InterPro:IPR008273), Phosphatidylinositol transfer protein-like, N-terminal (InterPro:IPR011074); BEST Arabidopsis thaliana protein match is: SEC14 cytosolic factor family protein / phosphoglyceride transfer family protein (TAIR:AT4G08690.2); Has 1880 Blast hits to 1880 proteins in 171 species: Archae - 0; Bacteria - 0; Metazoa - 814; Fungi - 399; Plants - 444; Viruses - 0; Other Eukaryotes - 223 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.1654.1.S1_s_at | 0.0336887 | 7.77462 | Glyma15g01690.2 | AT1G52360.1 | coatomer protein complex, subunit beta 2 (beta prime), putative; FUNCTIONS IN: protein binding, structural molecule activity; INVOLVED IN: intracellular protein transport, vesicle-mediated transport; LOCATED IN: membrane coat, COPI vesicle coat; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Coatomer, WD associated region (InterPro:IPR006692), Cytochrome cd1-nitrite reductase-like, C-terminal haem d1 (InterPro:IPR011048), WD40 repeat (InterPro:IPR001680), WD40 repeat-like (InterPro:IPR011046), WD40 repeat, region (InterPro:IPR017986), Coatomer, beta' subunit (InterPro:IPR016453), WD40/YVTN repeat-like (InterPro:IPR015943); BEST Arabidopsis thaliana protein match is: coatomer protein complex, subunit beta 2 (beta prime), putative (TAIR:AT3G15980.3); Has 57972 Blast hits to 24958 proteins in 644 species: Archae - 48; Bacteria - 5666; Metazoa - 27314; Fungi - 11164; Plants - 5401; Viruses - 42; Other Eukaryotes - 8337 (source: NCBI BLink). |
| Gma.7677.1.A1_at | 0.0999979 | 4.90152 | Glyma15g22200.1 | AT4G19490.2 | Putative homolog of yeast Vps54. Thought to associate with POK and ATVPS53 in a plant GARP-like complex involved in the membrane trafficking system. |
| GmaAffx.65026.1.S1_at | 0.0786998 | 3.23859 | Glyma07g37790.1 | AT1G09070.1 | SRC2 specifically binds the peptide PIEPPPHH, and moves from ER to a vacuole fraction where it gets internalized. Involved in Protein Storage Vacuole targeting. |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.3108.2.S1_at | 0.0267886 | 2.47486 | Glyma01g31600.1 | AT2G38360.1 | PRENYLATED RAB ACCEPTOR 1.B4 (PRA1.B4); FUNCTIONS IN: molecular_function unknown; INVOLVED IN: vesicle-mediated transport; LOCATED IN: endoplasmic reticulum; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Prenylated rab acceptor PRA1 (InterPro:IPR004895); BEST Arabidopsis thaliana protein match is: PRA1.B1 (PRENYLATED RAB ACCEPTOR 1.B1) (TAIR:AT3G56110.1); Has 342 Blast hits to 342 proteins in 90 species: Archae - 0; Bacteria - 0; Metazoa - 100; Fungi - 35; Plants - 176; Viruses - 0; Other Eukaryotes - 31 (source: NCBI BLink). |
| GmaAffx.8297.1.S1_at | 0.0317164 | 2.36764 | Glyma19g40160.1 | AT3G05710.2 | member of SYP4 Gene Family |
| GmaAffx.3108.1.S1_at | 0.0765248 | 2.16707 | Glyma03g05630.1 | AT2G38360.1 | PRENYLATED RAB ACCEPTOR 1.B4 (PRA1.B4); FUNCTIONS IN: molecular_function unknown; INVOLVED IN: vesicle-mediated transport; LOCATED IN: endoplasmic reticulum; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Prenylated rab acceptor PRA1 (InterPro:IPR004895); BEST Arabidopsis thaliana protein match is: PRA1.B1 (PRENYLATED RAB ACCEPTOR 1.B1) (TAIR:AT3G56110.1); Has 342 Blast hits to 342 proteins in 90 species: Archae - 0; Bacteria - 0; Metazoa - 100; Fungi - 35; Plants - 176; Viruses - 0; Other Eukaryotes - 31 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.17743.1.A1_at | 0.0598831 | 1.9744 | No Soybean Match found | AT2G06530.1 | VPS2.1; INVOLVED IN: vesicle-mediated transport; LOCATED IN: ESCRT III complex; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Snf7 (InterPro:IPR005024); BEST Arabidopsis thaliana protein match is: VPS2.3 (VACUOLAR PROTEIN SORTING-ASSOCIATED PROTEIN 2.3) (TAIR:AT1G03950.1); Has 1847 Blast hits to 1844 proteins in 191 species: Archae - 6; Bacteria - 15; Metazoa - 919; Fungi - 326; Plants - 342; Viruses - 6; Other Eukaryotes - 233 (source: NCBI BLink). |
| Gma.2608.1.S1_at | 0.0721004 | 1.84508 | Glyma17g14020.1 | AT4G11740.1 | Isolated as a suppressor of a dominant mutant in the Ara4 gene that was expressed in yeast ypt1 mutant strains. A novel protein with a small region of similarity to coil-coiled domain of yeast VSP27 protein. |
| Gma.16952.1.A1_at | 0.0272696 | 1.82225 | Glyma16g27050.1 | AT5G08080.1 | member of SYP13 Gene Family |
| GmaAffx.84502.1.S1_s_at | 0.07755 | 1.80849 | Glyma06g11710.1 | AT1G21900.1 | emp24/gp25L/p24 family protein; FUNCTIONS IN: protein transmembrane transporter activity; INVOLVED IN: intracellular protein transport, transport; LOCATED IN: endomembrane system, integral to membrane, membrane; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: GOLD (InterPro:IPR009038), emp24/gp25L/p24 (InterPro:IPR000348); BEST Arabidopsis thaliana protein match is: emp24/gp25L/p24 family protein (TAIR:AT1G09580.1); Has 1198 Blast hits to 1196 proteins in 170 species: Archae - 0; Bacteria - 0; Metazoa - 609; Fungi - 316; Plants - 130; Viruses - 0; Other Eukaryotes - 143 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.4131.2.S1_at | 0.0721004 | 1.74436 | Ambiguous Hit | AT2G27600.1 | Encodes a SKD1 (Suppressor of K+ Transport Growth Defect1) homolog. Localized to the cytoplasm and to multivesicular endosomes. Involved in multivesicular endosome function. |
| GmaAffx.71118.1.A1_s_at | 0.0341303 | 1.73043 | Glyma06g01120.3 | AT4G34660.1 | SH3 domain-containing protein 2 (SH3P2); FUNCTIONS IN: clathrin binding; INVOLVED IN: biological_process unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Src homology-3 domain (InterPro:IPR001452); BEST Arabidopsis thaliana protein match is: clathrin binding (TAIR:AT4G18060.1); Has 1201 Blast hits to 1169 proteins in 144 species: Archae - 0; Bacteria - 12; Metazoa - 956; Fungi - 47; Plants - 87; Viruses - 0; Other Eukaryotes - 99 (source: NCBI BLink). |
| GmaAffx.3108.3.S1_at | 0.0998527 | 1.68306 | Glyma03g05630.1 | AT2G38360.1 | PRENYLATED RAB ACCEPTOR 1.B4 (PRA1.B4); FUNCTIONS IN: molecular_function unknown; INVOLVED IN: vesicle-mediated transport; LOCATED IN: endoplasmic reticulum; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Prenylated rab acceptor PRA1 (InterPro:IPR004895); BEST Arabidopsis thaliana protein match is: PRA1.B1 (PRENYLATED RAB ACCEPTOR 1.B1) (TAIR:AT3G56110.1); Has 342 Blast hits to 342 proteins in 90 species: Archae - 0; Bacteria - 0; Metazoa - 100; Fungi - 35; Plants - 176; Viruses - 0; Other Eukaryotes - 31 (source: NCBI BLink). |
| GmaAffx.4493.1.A1_s_at | 0.0955455 | 1.61649 | Glyma01g33380.2 | AT2G45200.1 | Encodes a member of the GOS1 (Golgi SNARE) gene family. |
| | | | | | |

| | | | | | |
|---|---|---|---|---|---|
| Gma.11292.1.S1_s_at | 0.0681843 | 1.54167 | Glyma05g05260.2 | AT4G17530.1 | ATRAB1C; FUNCTIONS IN: GTP binding; INVOLVED IN: protein transport, small GTPase mediated signal transduction; LOCATED IN: plasma membrane, vacuole, membrane; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Ras GTPase (InterPro:IPR001806), Small GTP-binding protein (InterPro:IPR005225), Ras (InterPro:IPR013753), Ras small GTPase, Rab type (InterPro:IPR003579); BEST Arabidopsis thaliana protein match is: ATRAB1A; GTP binding (TAIR:AT5G47200.1); Has 23593 Blast hits to 23542 proteins in 647 species: Archae - 17; Bacteria - 111; Metazoa - 13220; Fungi - 2827; Plants - 2160; Viruses - 19; Other Eukaryotes - 5239 (source: NCBI BLink). |
| Gma.2787.1.S1_s_at | 0.0755306 | 1.54027 | Glyma09g00430.1 | AT1G59610.1 | A high molecular weight GTPase whose GTP-binding domain shows a low homology to those of other plant dynamin-like proteins.Contains a pleckstrin homologous domain. DRP2B and DRP1A participate together in clathrin-coated vesicle formation during endocytosis. |
| Gma.3590.1.S1_at | 0.0500445 | 1.52878 | Ambiguous Hit | AT1G15130.1 | hydroxyproline-rich glycoprotein family protein; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: BRO1 (InterPro:IPR004328); Has 24106 Blast hits to 13529 proteins in 779 species: Archae - 27; Bacteria - 1802; Metazoa - 9356; Fungi - 3772; Plants - 5606; Viruses - 598; Other Eukaryotes - 2945 (source: NCBI BLink). |

| Gma.3977.1.S1_s_at | 0.07755 | 1.49271 | Glyma18g02590.1 | AT3G56190.1 | Encodes one of two alpha-SNAPs (soluble NSF attachment protein) in Arabidopsis |
|---|---|---|---|---|---|
| GmaAffx.41532.1.S1_at | 0.0581114 | 1.45239 | Glyma13g24990.1 | AT2G20760.1 | protein binding / structural molecule; FUNCTIONS IN: protein binding, structural molecule activity; INVOLVED IN: intracellular protein transport, vesicle-mediated transport; LOCATED IN: plasma membrane; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Clathrin light chain (InterPro:IPR000996); BEST Arabidopsis thaliana protein match is: protein binding / structural molecule (TAIR:AT3G51890.1); Has 1552 Blast hits to 1056 proteins in 209 species: Archae - 0; Bacteria - 461; Metazoa - 500; Fungi - 90; Plants - 101; Viruses - 0; Other Eukaryotes - 400 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.8341.1.S1_at | 0.0990256 | 1.38028 | Glyma02g39720.1 | AT1G07030.1 | mitochondrial substrate carrier family protein; FUNCTIONS IN: binding; INVOLVED IN: transport, mitochondrial transport; LOCATED IN: mitochondrial inner membrane, membrane; EXPRESSED IN: 20 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Mitochondrial carrier protein (InterPro:IPR002067), Mitochondrial substrate carrier (InterPro:IPR001993), Mitochondrial substrate/solute carrier (InterPro:IPR018108); BEST Arabidopsis thaliana protein match is: mitochondrial substrate carrier family protein (TAIR:AT2G30160.1); Has 19774 Blast hits to 10070 proteins in 353 species: Archae - 0; Bacteria - 0; Metazoa - 9836; Fungi - 5154; Plants - 2928; Viruses - 0; Other Eukaryotes - 1856 (source: NCBI BLink). |
| Gma.6058.1.A1_at | 0.0576293 | 1.35865 | Glyma10g33960.1 | AT5G04850.1 | VPS60.2; INVOLVED IN: vesicle-mediated transport; LOCATED IN: cellular_component unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Snf7 (InterPro:IPR005024); BEST Arabidopsis thaliana protein match is: VPS60.1 (TAIR:AT3G10640.1); Has 906 Blast hits to 906 proteins in 160 species: Archae - 3; Bacteria - 0; Metazoa - 432; Fungi - 204; Plants - 144; Viruses - 0; Other Eukaryotes - 123 (source: NCBI BLink). |
| Gma.1152.1.S1_a_at | 0.0687604 | 1.3101 | Glyma03g03120.2 | AT1G71820.1 | Encodes a member of the exocyst complex gene family. The exocyst is a protein complex involved in tethering vesicles to the plasma membrane during regulated or polarized secretion. |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.8845.2.S1_at | 0.0916528 | -4.0956 | Glyma20g00730.1 | AT2G39970.1 | peroxisomal membrane protein (PMP36); FUNCTIONS IN: binding; INVOLVED IN: transport; LOCATED IN: peroxisomal membrane, mitochondrial inner membrane, plasma membrane; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Mitochondrial substrate carrier (InterPro:IPR001993), Mitochondrial substrate/solute carrier (InterPro:IPR018108); BEST Arabidopsis thaliana protein match is: ATFOLT1 (ARABIDOPSIS THALIANA FOLATE TRANSPORTER 1); binding / folic acid transporter (TAIR:AT5G66380.1); Has 12085 Blast hits to 8258 proteins in 311 species: Archae - 0; Bacteria - 0; Metazoa - 6388; Fungi - 3049; Plants - 1606; Viruses - 0; Other Eukaryotes - 1042 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.8307.2.A1_at | 0.0862961 | -2.2418 | Glyma15g12730.1 | AT5G47730.1 | SEC14 cytosolic factor, putative / polyphosphoinositide-binding protein, putative; FUNCTIONS IN: transporter activity, binding; INVOLVED IN: transport; LOCATED IN: cellular_component unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Cellular retinaldehyde-binding/triple function, C-terminal (InterPro:IPR001251), Cellular retinaldehyde-binding/triple function, N-terminal (InterPro:IPR008273), Phosphatidylinositol transfer protein-like, N-terminal (InterPro:IPR011074); BEST Arabidopsis thaliana protein match is: SEC14 cytosolic factor (SEC14) / phosphoglyceride transfer protein (TAIR:AT1G55840.1); Has 1254 Blast hits to 1254 proteins in 165 species: Archae - 0; Bacteria - 0; Metazoa - 467; Fungi - 271; Plants - 377; Viruses - 0; Other Eukaryotes - 139 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.77471.1.S1_at | 0.0537629 | -1.8174 | Glyma20g37500.1 | AT5G27395.2 | P-P-bond-hydrolysis-driven protein transmembrane transporter; FUNCTIONS IN: P-P-bond-hydrolysis-driven protein transmembrane transporter activity; INVOLVED IN: intracellular protein transport; LOCATED IN: mitochondrion, mitochondrial inner membrane presequence translocase complex; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Mitochondrial import inner membrane translocase, subunit Tim44 (InterPro:IPR007379); Has 249 Blast hits to 249 proteins in 90 species: Archae - 0; Bacteria - 92; Metazoa - 76; Fungi - 0; Plants - 17; Viruses - 0; Other Eukaryotes - 64 (source: NCBI BLink). |
| Gma.10328.1.S1_x_at | 0.0758646 | -1.2299 | Glyma15g01660.2 | AT1G52380.1 | Ran-binding protein 1 domain-containing protein / RanBP1 domain-containing protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: intracellular transport; LOCATED IN: cytosol, nucleus; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Ran Binding Protein 1 (InterPro:IPR000156), Pleckstrin homology-type (InterPro:IPR011993), NUP50 (Nucleoporin 50 kDa) (InterPro:IPR015007); BEST Arabidopsis thaliana protein match is: Ran-binding protein 1 domain-containing protein / RanBP1 domain-containing protein (TAIR:AT3G15970.1); Has 1224 Blast hits to 974 proteins in 166 species: Archae - 0; Bacteria - 22; Metazoa - 715; Fungi - 281; Plants - 101; Viruses - 2; Other Eukaryotes - 103 (source: NCBI BLink). |

Table 3 Candidate genes related to metabolism

| Affy ID | Q-value | Fold change | Soybean Gene | Arabidopsis homolog | Description |
|---|---|---|---|---|---|
| GmaAffx.54955.1.S1_ | 0.02727 | 24.314 | No Soybean Match Identified | AT5G12210.2 | geranylgeranyl transferase type II beta subunit, putative / RAB geranylgeranyltransferase beta subunit, putative; FUNCTIONS IN: catalytic activity; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Terpenoid cylases/protein prenyltransferase alpha-alpha toroid (InterPro:IPR008930), Prenyltransferase/squalene oxidase (InterPro:IPR001330); BEST Arabidopsis thaliana protein match is: geranylgeranyl transferase type II beta subunit, putative / RAB geranylgeranyltransferase beta subunit, putative (TAIR:AT3G12070.2); Has 1176 Blast hits to 1014 proteins in 185 species: Archae - 0; Bacteria - 22; Metazoa - 490; Fungi - 275; Plants - 123; Viruses - 0; Other Eukaryotes - 266 (source: NCBI BLink). |
| GmaAffx.20567.1.S1_ | 0.05952 | 12.618 | Glyma06g03420.1 | AT1G19530.1 | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: N-terminal protein myristoylation; LOCATED IN: cellular_component unknown; EXPRESSED IN: inflorescence meristem, leaf apex, hypocotyl, flower, root; EXPRESSED DURING: petal differentiation and expansion stage; Has 28 Blast hits to 28 proteins in 7 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 28; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| Gma.9.1.S1_at | 0.06264 | 11.495 | Glyma04g01190.2 | AT4G39140.5 | protein binding / zinc ion binding; FUNCTIONS IN: protein binding, zinc ion binding; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Zinc finger, RING-type (InterPro:IPR001841); BEST Arabidopsis thaliana protein match is: protein binding / zinc ion binding (TAIR:AT2G21500.2); Has 1494 Blast hits to 203 proteins in 51 species: Archae - 0; Bacteria - 6; Metazoa - 85; Fungi - 41; Plants - 92; Viruses - 0; Other Eukaryotes - 1270 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.7861.1.S1_at | 0.08422 | 11.452 | Glyma04g08020.1 | AT5G10695.1 | unknown protein; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 12 growth stages; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT5G57123.1); Has 54 Blast hits to 54 proteins in 8 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 54; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| Gma.9069.1.S1_at | 0.09487 | 6.8757 | Glyma14g02630.1 | AT4G38460.1 | geranylgeranyl reductase (GGR); FUNCTIONS IN: farnesyltranstransferase activity; INVOLVED IN: isoprenoid biosynthetic process; LOCATED IN: chloroplast; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Polyprenyl synthetase-related (InterPro:IPR017446), Terpenoid synthase (InterPro:IPR008949), Polyprenyl synthetase (InterPro:IPR000092); BEST Arabidopsis thaliana protein match is: GGPS1 (GERANYLGERANYL PYROPHOSPHATE SYNTHASE 1); farnesyltranstransferase (TAIR:AT4G36810.1); Has 9804 Blast hits to 9803 proteins in 1646 species: Archae - 213; Bacteria - 4792; Metazoa - 110; Fungi - 109; Plants - 266; Viruses - 6; Other Eukaryotes - 4308 (source: NCBI BLink). |
| GmaAffx.11201.1.A1 | 0.02679 | 6.8076 | Glyma19g42560.1 | AT1G54790.1 | GDSL-motif lipase/hydrolase family protein; FUNCTIONS IN: hydrolase activity, acting on ester bonds, carboxylesterase activity; INVOLVED IN: lipid metabolic process; LOCATED IN: endomembrane system; EXPRESSED IN: 16 plant structures; EXPRESSED DURING: 8 growth stages; CONTAINS InterPro DOMAIN/s: Lipase, GDSL (InterPro:IPR001087); BEST Arabidopsis thaliana protein match is: GDSL-motif lipase/hydrolase family protein (TAIR:AT3G05180.1); Has 1592 Blast hits to 1570 proteins in 87 species: Archae - 0; Bacteria - 61; Metazoa - 1; Fungi - 4; Plants - 1524; Viruses - 0; Other Eukaryotes - 2 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.9838.1.S1_at | 0.01639 | 6.649 | Glyma01g37550.1 | AT3G51030.1 | encodes a cytosolic thioredoxin that reduces disulfide bridges of target proteins by the reversible formation of a disulfide bridge between two neighboring Cys residues present in the active site. Thioredoxins have been found to regulate a variety of biological reactions in prokaryotic and eukaryotic cells. |
| GmaAffx.87290.1.S1_ | 0.09399 | 6.2412 | Glyma10g31540.2 | AT1G32900.1 | starch synthase, putative; FUNCTIONS IN: transferase activity, transferring glycosyl groups; INVOLVED IN: biosynthetic process, glucan biosynthetic process; LOCATED IN: chloroplast; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Glycogen/starch synthases, ADP-glucose type (InterPro:IPR011835), Starch synthase catalytic region (InterPro:IPR013534), Glycosyl transferase, group 1 (InterPro:IPR001296); BEST Arabidopsis thaliana protein match is: AtSS2 (starch synthase 2); transferase, transferring glycosyl groups (TAIR:AT3G01180.1); Has 9548 Blast hits to 9532 proteins in 2484 species: Archae - 212; Bacteria - 3723; Metazoa - 12; Fungi - 137; Plants - 4346; Viruses - 0; Other Eukaryotes - 1118 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.10412.1.A1_at | 0.04364 | 5.917 | Glyma01g42440.1 | AT1G63000.1 | NUCLEOTIDE-RHAMNOSE SYNTHASE/EPIMERASE-REDUCTASE (NRS/ER); FUNCTIONS IN: UDP-4-keto-rhamnose-4-keto-reductase activity, dTDP-4-dehydrorhamnose reductase activity, UDP-4-keto-6-deoxy-glucose-3,5-epimerase activity, dTDP-4-dehydrorhamnose 3,5-epimerase activity; INVOLVED IN: dTDP-rhamnose biosynthetic process, UDP-rhamnose biosynthetic process; LOCATED IN: soluble fraction, plasma membrane; EXPRESSED IN: 26 plant structures; EXPRESSED DURING: 16 growth stages; CONTAINS InterPro DOMAIN/s: NAD-dependent epimerase/dehydratase (InterPro:IPR001509), NAD(P)-binding (InterPro:IPR016040); BEST Arabidopsis thaliana protein match is: RHM1 (RHAMNOSE BIOSYNTHESIS 1); UDP-L-rhamnose synthase/ UDP-glucose 4,6-dehydratase/ catalytic (TAIR:AT1G78570.1); Has 882 Blast hits to 882 proteins in 235 species: Archae - 35; Bacteria - 346; Metazoa - 9; Fungi - 16; Plants - 160; Viruses - 5; Other Eukaryotes - 311 (source: NCBI BLink). |
| GmaAffx.14279.1.S1_ | 0.07755 | 5.4498 | Glyma19g40550.1 | AT4G01970.1 | Encodes a putative stachyose synthetase. |
| GmaAffx.494.1.S1_at | 0.02679 | 4.9898 | Glyma16g28970.2 | AT5G24090.1 | Chitinase A (class III) expressed exclusively under environmental stress conditions. |

| | | | | | |
|---|---|---|---|---|---|
| Gma.939.1.A1_at | 0.07652 | 4.9681 | Glyma07g38790.1 | AT1G54870.1 | binding / catalytic/ oxidoreductase; FUNCTIONS IN: oxidoreductase activity, binding, catalytic activity; INVOLVED IN: metabolic process; LOCATED IN: chloroplast; EXPRESSED IN: 8 plant structures; EXPRESSED DURING: M germinated pollen stage, 4 anthesis, petal differentiation and expansion stage; CONTAINS InterPro DOMAIN/s: NAD(P)-binding (InterPro:IPR016040), Glucose/ribitol dehydrogenase (InterPro:IPR002347), Short-chain dehydrogenase/reductase SDR (InterPro:IPR002198); BEST Arabidopsis thaliana protein match is: short-chain dehydrogenase/reductase (SDR) family protein (TAIR:AT3G05260.1); Has 83077 Blast hits to 82922 proteins in 2231 species: Archae - 482; Bacteria - 45268; Metazoa - 5002; Fungi - 4212; Plants - 1573; Viruses - 5; Other Eukaryotes - 26535 (source: NCBI BLink). |
| GmaAffx.40696.1.S1_ | 0.05948 | 4.9652 | Glyma01g43470.5 | AT4G23850.1 | long-chain-fatty-acid--CoA ligase / long-chain acyl-CoA synthetase; FUNCTIONS IN: catalytic activity; INVOLVED IN: fatty acid biosynthetic process; LOCATED IN: plasma membrane; EXPRESSED IN: 26 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: AMP-dependent synthetase and ligase (InterPro:IPR000873); BEST Arabidopsis thaliana protein match is: long-chain-fatty-acid--CoA ligase, putative / long-chain acyl-CoA synthetase, putative (TAIR:AT4G11030.1); Has 34929 Blast hits to 33173 proteins in 1967 species: Archae - 465; Bacteria - 18245; Metazoa - 1999; Fungi - 1232; Plants - 1153; Viruses - 1; Other Eukaryotes - 11834 (source: NCBI BLink). |
| GmaAffx.84164.1.A1 | 0.0472 | 4.623 | Glyma08g11330.1 | AT4G15550.1 | UDP-glucose:indole-3-acetate beta-D-glucosyltransferase |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.19618.1.A1 | 0.05458 | 3.9626 | Glyma08g19430.1 | AT1G43670.1 | fructose-1,6-bisphosphatase, putative / D-fructose-1,6-bisphosphate 1-phosphohydrolase, putative / FBPase, putative; FUNCTIONS IN: fructose 1,6-bisphosphate 1-phosphatase activity, phosphoric ester hydrolase activity; INVOLVED IN: carbohydrate metabolic process, fructose metabolic process; LOCATED IN: cellular_component unknown; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 16 growth stages; CONTAINS InterPro DOMAIN/s: Inositol monophosphatase/Fructose-1,6-bisphosphatase (InterPro:IPR017955), Fructose-1,6-bisphosphatase (InterPro:IPR000146); BEST Arabidopsis thaliana protein match is: fructose-1,6-bisphosphatase, putative / D-fructose-1,6-bisphosphate 1-phosphohydrolase, putative / FBPase, putative (TAIR:AT3G54050.1); Has 2339 Blast hits to 2334 proteins in 798 species: Archae - 22; Bacteria - 1225; Metazoa - 341; Fungi - 107; Plants - 205; Viruses - 0; Other Eukaryotes - 439 (source: NCBI BLink). |
| GmaAffx.51101.1.S1_ | 0.05722 | 3.6461 | Glyma10g39060.1 | AT2G32020.1 | GCN5-related N-acetyltransferase (GNAT) family protein; FUNCTIONS IN: N-acetyltransferase activity; INVOLVED IN: response to abscisic acid stimulus, metabolic process; LOCATED IN: cellular_component unknown; EXPRESSED IN: stem, cotyledon, hypocotyl, root, stamen; EXPRESSED DURING: 4 anthesis, petal differentiation and expansion stage; CONTAINS InterPro DOMAIN/s: GCN5-related N-acetyltransferase (InterPro:IPR000182), Acyl-CoA N-acyltransferase (InterPro:IPR016181); BEST Arabidopsis thaliana protein match is: GCN5-related N-acetyltransferase (GNAT) family protein (TAIR:AT2G32030.1); Has 5194 Blast hits to 5189 proteins in 878 species: Archae - 65; Bacteria - 4464; Metazoa - 1; Fungi - 93; Plants - 71; Viruses - 2; Other Eukaryotes - 498 (source: NCBI BLink). |
| Gma.6802.1.A1_at | 0.05952 | 3.5908 | Glyma19g45230.1 | AT1G53920.1 | Contains lipase signature motif and GDSL domain. |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.91156.1.S1_ | 0.06018 | 3.5206 | Glyma14g38890.1 | AT3G46940.1 | deoxyuridine 5'-triphosphate nucleotidohydrolase family; FUNCTIONS IN: hydrolase activity, dUTP diphosphatase activity; INVOLVED IN: 2'-deoxyribonucleotide metabolic process, dUTP metabolic process; LOCATED IN: cellular_component unknown; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: DeoxyUTP pyrophosphatase subfamily 1 (InterPro:IPR008181), DeoxyUTP pyrophosphatase (InterPro:IPR008180); Has 5469 Blast hits to 5455 proteins in 1432 species: Archae - 13; Bacteria - 2270; Metazoa - 201; Fungi - 106; Plants - 30; Viruses - 721; Other Eukaryotes - 2128 (source: NCBI BLink). |
| GmaAffx.41357.2.S1_ | 0.0448 | 3.4125 | Glyma02g34640.1 | AT2G41490.1 | UDP-GlcNAc:dolichol phosphate N-acetylglucosamine-1-phosphate transferase |
| GmaAffx.50381.1.S1_ | 0.05603 | 3.124 | Glyma05g34670.1 | AT2G01670.1 | Arabidopsis thaliana Nudix hydrolase homolog 17 (atnudt17); FUNCTIONS IN: hydrolase activity; INVOLVED IN: biological_process unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: NUDIX (InterPro:IPR015797), NUDIX hydrolase, core (InterPro:IPR000086); BEST Arabidopsis thaliana protein match is: atnudt18 (Arabidopsis thaliana Nudix hydrolase homolog 18); hydrolase (TAIR:AT1G14860.1); Has 671 Blast hits to 669 proteins in 184 species: Archae - 0; Bacteria - 183; Metazoa - 198; Fungi - 77; Plants - 132; Viruses - 0; Other Eukaryotes - 81 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.67906.1.S1_ | 0.05376 | 3.0555 | Glyma18g04090.1 | AT3G55550.1 | lectin protein kinase, putative; FUNCTIONS IN: kinase activity; INVOLVED IN: protein amino acid phosphorylation; LOCATED IN: plasma membrane; CONTAINS InterPro DOMAIN/s: Legume lectin, beta domain (InterPro:IPR001220), Protein kinase, ATP binding site (InterPro:IPR017441), Serine/threonine protein kinase-related (InterPro:IPR017442), Concanavalin A-like lectin/glucanase, subgroup (InterPro:IPR013320), Protein kinase-like (InterPro:IPR011009), Serine/threonine protein kinase, active site (InterPro:IPR008271), Protein kinase, core (InterPro:IPR000719), Concanavalin A-like lectin/glucanase (InterPro:IPR008985); BEST Arabidopsis thaliana protein match is: RLK (receptor lectin kinase); kinase (TAIR:AT2G37710.1); Has 86229 Blast hits to 85142 proteins in 3260 species: Archae - 47; Bacteria - 7439; Metazoa - 37868; Fungi - 6742; Plants - 19531; Viruses - 366; Other Eukaryotes - 14236 (source: NCBI BLink). |
| Gma.7147.1.A1_at | 0.02316 | 2.9824 | Glyma09g34110.1 | AT1G32200.2 | Encodes a chloroplast glycerol-3-phosphate acyltransferase.Involved in the biosynthesis of chloroplast phosphatidylglycerol. |
| GmaAffx.14279.2.A1 | 0.07586 | 2.9728 | Glyma19g40550.1 | AT4G01970.1 | Encodes a putative stachyose synthetase. |
| Gma.10242.1.A1_at | 0.0951 | 2.907 | Glyma07g33490.1 | AT3G19240.1 | FUNCTIONS IN: molecular_function unknown; INVOLVED IN: N-terminal protein myristoylation; LOCATED IN: cellular_component unknown; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: WD40 repeat-like (InterPro:IPR011046), Vacuolar import and degradation, Vid27-related (InterPro:IPR013863); BEST Arabidopsis thaliana protein match is: dem protein-related / defective embryo and meristems protein-related (TAIR:AT4G33400.1); Has 206 Blast hits to 202 proteins in 88 species: Archae - 0; Bacteria - 0; Metazoa - 11; Fungi - 117; Plants - 38; Viruses - 0; Other Eukaryotes - 40 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.10581.1.S1_at | 0.05376 | 2.8286 | Glyma08g20820.1 | AT1G17710.1 | phosphatase; FUNCTIONS IN: phosphatase activity; INVOLVED IN: metabolic process; EXPRESSED IN: 9 plant structures; EXPRESSED DURING: L mature pollen stage, M germinated pollen stage, 4 anthesis, petal differentiation and expansion stage; CONTAINS InterPro DOMAIN/s: Pyridoxal phosphate phosphatase, PHOSPHO2 (InterPro:IPR016965), HAD superfamily hydrolase, subfamily IB, PSPase-like (InterPro:IPR006383), Pyridoxal phosphate phosphatase-related (InterPro:IPR006384); BEST Arabidopsis thaliana protein match is: phosphatase (TAIR:AT1G73010.1); Has 262 Blast hits to 260 proteins in 77 species: Archae - 0; Bacteria - 12; Metazoa - 155; Fungi - 12; Plants - 57; Viruses - 0; Other Eukaryotes - 26 (source: NCBI BLink). |
| GmaAffx.20724.1.S1_ | 0.06187 | 2.7918 | Glyma10g35700.1 | AT5G60390.3 | elongation factor 1-alpha / EF-1-alpha; FUNCTIONS IN: calmodulin binding, translation elongation factor activity; INVOLVED IN: translational elongation; LOCATED IN: mitochondrion, nucleus, cytoplasm; EXPRESSED IN: cotyledon, male gametophyte, guard cell, pollen tube, seed; EXPRESSED DURING: L mature pollen stage, M germinated pollen stage; CONTAINS InterPro DOMAIN/s: Translation elongation factor EFTu/EF1A, C-terminal (InterPro:IPR004160), Translation elongation factor EFTu/EF1A, domain 2 (InterPro:IPR004161), Translation elongation factor EF1A/initiation factor IF2gamma, C-terminal (InterPro:IPR009001), Protein synthesis factor, GTP-binding (InterPro:IPR000795), Translation elongation and initiation factors/Ribosomal, beta-barrel (InterPro:IPR009000), Translation elongation factor EF1A, eukaryotic and archaeal (InterPro:IPR004539); BEST Arabidopsis thaliana protein match is: elongation factor 1-alpha / EF-1-alpha (TAIR:AT1G07940.2); Has 56847 Blast hits to 56774 proteins in 13741 species: Archae - 656; Bacteria - 18823; Metazoa - 15489; Fungi - 8699; Plants - 1228; Viruses - 3; Other Eukaryotes - 11949 (source: NCBI BLink |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.82961.1.S1_ | 0.06865 | 2.7523 | Glyma20g25550.3 | AT2G01220.2 | nucleotidyltransferase; FUNCTIONS IN: nucleotidyltransferase activity; INVOLVED IN: biosynthetic process; LOCATED IN: chloroplast envelope; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Rossmann-like alpha/beta/alpha sandwich fold (InterPro:IPR014729), Cytidylyltransferase (InterPro:IPR004820); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT3G27610.1); Has 373 Blast hits to 367 proteins in 37 species: Archae - 6; Bacteria - 25; Metazoa - 0; Fungi - 5; Plants - 35; Viruses - 0; Other Eukaryotes - 302 (source: NCBI BLink). |
| GmaAffx.12665.1.A1_ | 0.02679 | 2.72 | Glyma03g41580.1 | AT4G01130.1 | acetylesterase, putative; FUNCTIONS IN: hydrolase activity, acting on ester bonds, carboxylesterase activity; INVOLVED IN: lipid metabolic process; LOCATED IN: endomembrane system; EXPRESSED IN: 18 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Lipase, GDSL (InterPro:IPR001087); BEST Arabidopsis thaliana protein match is: GDSL-motif lipase/hydrolase family protein (TAIR:AT3G26430.1); Has 1707 Blast hits to 1677 proteins in 112 species: Archae - 0; Bacteria - 95; Metazoa - 1; Fungi - 41; Plants - 1566; Viruses - 0; Other Eukaryotes - 4 (source: NCBI BLink). |
| GmaAffx.47486.1.S1_ | 0.09221 | 2.6939 | Glyma14g40490.1 | AT1G75690.1 | chaperone protein dnaJ-related; FUNCTIONS IN: unfolded protein binding, heat shock protein binding; INVOLVED IN: protein folding; LOCATED IN: chloroplast thylakoid membrane, chloroplast stroma, chloroplast; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Heat shock protein DnaJ, cysteine-rich region (InterPro:IPR001305); Has 326 Blast hits to 308 proteins in 90 species: Archae - 4; Bacteria - 115; Metazoa - 13; Fungi - 12; Plants - 104; Viruses - 0; Other Eukaryotes - 78 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.440.1.S1_at | 0.02316 | 2.6748 | Glyma07g17240.1 | AT2G38000.1 | chaperone protein dnaJ-related; LOCATED IN: cellular_component unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 15 growth stages; Has 643 Blast hits to 619 proteins in 228 species: Archae - 19; Bacteria - 379; Metazoa - 79; Fungi - 4; Plants - 29; Viruses - 0; Other Eukaryotes - 133 (source: NCBI BLink). |
| GmaAffx.93457.1.S1_ | 0.08352 | 2.647 | Glyma18g53170.1 | AT5G38200.1 | hydrolase; FUNCTIONS IN: hydrolase activity; INVOLVED IN: glutamine metabolic process; LOCATED IN: cellular_component unknown; CONTAINS InterPro DOMAIN/s: Peptidase C26 (InterPro:IPR011697), Glutamine amidotransferase type 1 (InterPro:IPR017926); BEST Arabidopsis thaliana protein match is: hydrolase (TAIR:AT1G66860.1); Has 2060 Blast hits to 2059 proteins in 580 species: Archae - 8; Bacteria - 1443; Metazoa - 0; Fungi - 2; Plants - 26; Viruses - 0; Other Eukaryotes - 581 (source: NCBI BLink). |
| Gma.1583.1.S1_s_at | 0.09403 | 2.6449 | Glyma02g01720.1 | AT1G02660.1 | lipase class 3 family protein; FUNCTIONS IN: triacylglycerol lipase activity; INVOLVED IN: lipid metabolic process; LOCATED IN: cellular_component unknown; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Lipase, class 3 (InterPro:IPR002921); BEST Arabidopsis thaliana protein match is: lipase class 3 family protein (TAIR:AT3G62590.1); Has 601 Blast hits to 595 proteins in 116 species: Archae - 0; Bacteria - 17; Metazoa - 188; Fungi - 136; Plants - 92; Viruses - 14; Other Eukaryotes - 154 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.30467.1.S1_ | 0.03172 | 2.6436 | Glyma05g30870.2 | AT1G12990.1 | glycosyl transferase family 17 protein; FUNCTIONS IN: transferase activity, transferring glycosyl groups, acetylglucosaminyltransferase activity; INVOLVED IN: protein amino acid N-linked glycosylation; LOCATED IN: membrane; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Glycosyl transferase, family 17 (InterPro:IPR006813); BEST Arabidopsis thaliana protein match is: glycosyl transferase family 17 protein (TAIR:AT1G67880.1); Has 972 Blast hits to 971 proteins in 58 species: Archae - 0; Bacteria - 24; Metazoa - 46; Fungi - 23; Plants - 68; Viruses - 4; Other Eukaryotes - 807 (source: NCBI BLink). |
| Gma.4920.1.A1_at | 0.02679 | 2.5472 | Glyma14g02350.1 | AT5G55180.1 | glycosyl hydrolase family 17 protein; FUNCTIONS IN: cation binding, hydrolase activity, hydrolyzing O-glycosyl compounds, catalytic activity; INVOLVED IN: carbohydrate metabolic process; LOCATED IN: endomembrane system; EXPRESSED IN: 18 plant structures; EXPRESSED DURING: 11 growth stages; CONTAINS InterPro DOMAIN/s: X8 (InterPro:IPR012946), Glycoside hydrolase, catalytic core (InterPro:IPR017853), Glycoside hydrolase, family 17 (InterPro:IPR000490), Glycoside hydrolase, subgroup, catalytic core (InterPro:IPR013781); BEST Arabidopsis thaliana protein match is: catalytic/ cation binding / hydrolase, hydrolyzing O-glycosyl compounds (TAIR:AT4G26830.1); Has 1770 Blast hits to 1712 proteins in 115 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 8; Plants - 1757; Viruses - 0; Other Eukaryotes - 5 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.14475.1.S1_ | 0.03894 | 2.435 | Glyma14g37770.1 | AT2G30150.1 | UDP-glucoronosyl/UDP-glucosyl transferase family protein; FUNCTIONS IN: transferase activity, transferring hexosyl groups, UDP-glycosyltransferase activity, transferase activity, transferring glycosyl groups; INVOLVED IN: metabolic process; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: UDP-glucuronosyl/UDP-glucosyltransferase (InterPro:IPR002213); BEST Arabidopsis thaliana protein match is: UDP-glucoronosyl/UDP-glucosyl transferase family protein (TAIR:AT2G30140.1); Has 4165 Blast hits to 4145 proteins in 300 species: Archae - 0; Bacteria - 157; Metazoa - 1277; Fungi - 8; Plants - 2672; Viruses - 38; Other Eukaryotes - 13 (source: NCBI BLink). |
| Gma.5690.1.S1_at | 0.05105 | 2.4053 | Glyma10g42500.1 | AT3G54290.1 | FUNCTIONS IN: molecular_function unknown; INVOLVED IN: N-terminal protein myristoylation; LOCATED IN: chloroplast; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Haemerythrin/HHE cation-binding motif (InterPro:IPR012312); Has 35 Blast hits to 35 proteins in 7 species: Archae - 0; Bacteria - 2; Metazoa - 0; Fungi - 0; Plants - 32; Viruses - 0; Other Eukaryotes - 1 (source: NCBI BLink). |
| Gma.6847.1.A1_at | 0.09073 | 2.3783 | Glyma13g07230.1 | AT5G15870.1 | glycosyl hydrolase family 81 protein; FUNCTIONS IN: hydrolase activity, acting on glycosyl bonds, endo-1,3(4)-beta-glucanase activity; INVOLVED IN: response to cyclopentenone; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Glycoside hydrolase, family 81 (InterPro:IPR005200); BEST Arabidopsis thaliana protein match is: glycosyl hydrolase family 81 protein (TAIR:AT1G18310.1); Has 593 Blast hits to 503 proteins in 121 species: Archae - 0; Bacteria - 53; Metazoa - 84; Fungi - 150; Plants - 126; Viruses - 54; Other Eukaryotes - 126 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.19010.1.A1 | 0.07824 | 2.377 | Glyma10g29820.1 | AT1G54790.1 | GDSL-motif lipase/hydrolase family protein; FUNCTIONS IN: hydrolase activity, acting on ester bonds, carboxylesterase activity; INVOLVED IN: lipid metabolic process; LOCATED IN: endomembrane system; EXPRESSED IN: 16 plant structures; EXPRESSED DURING: 8 growth stages; CONTAINS InterPro DOMAIN/s: Lipase, GDSL (InterPro:IPR001087); BEST Arabidopsis thaliana protein match is: GDSL-motif lipase/hydrolase family protein (TAIR:AT3G05180.1); Has 1592 Blast hits to 1570 proteins in 87 species: Archae - 0; Bacteria - 61; Metazoa - 1; Fungi - 4; Plants - 1524; Viruses - 0; Other Eukaryotes - 2 (source: NCBI BLink). |
| Gma.6684.1.S1_at | 0.05811 | 2.3461 | Glyma18g01890.1 | AT5G65280.1 | Encodes a protein with reported similarity to GCR2 a putative G protein coupled receptor thought to be an ABA receptor. Loss of function mutations in GCL1 show no ABA response defects based on assays of seed germination and seedling development.GCL1 also has similarity to LANCL1 and LANCL2, human homologs of bacterial lanthionine synthetase. |
| Gma.3947.1.S1_at | 0.03172 | 2.3314 | Glyma03g33710.1 | AT5G03160.1 | J domain protein localized in ER lumen. Can partially compensate for the growth defect in jem1 scj1 mutant yeast. |
| GmaAffx.52033.1.S1_ | 0.0721 | 2.3293 | Glyma15g14040.1 | AT4G20840.1 | FAD-binding domain-containing protein; FUNCTIONS IN: electron carrier activity, oxidoreductase activity, FAD binding, catalytic activity; LOCATED IN: apoplast, plasma membrane; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 10 growth stages; CONTAINS InterPro DOMAIN/s: FAD-binding, type 2 (InterPro:IPR016166), Berberine/berberine-like (InterPro:IPR012951), FAD linked oxidase, N-terminal (InterPro:IPR006094); BEST Arabidopsis thaliana protein match is: FAD-binding domain-containing protein (TAIR:AT4G20830.2); Has 2748 Blast hits to 2654 proteins in 455 species: Archae - 10; Bacteria - 1082; Metazoa - 3; Fungi - 1106; Plants - 397; Viruses - 0; Other Eukaryotes - 150 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.92936.1.S1_ | 0.0819 | 2.3082 | Glyma07g17240.1 | AT2G38000.1 | chaperone protein dnaJ-related; LOCATED IN: cellular_component unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 15 growth stages; Has 643 Blast hits to 619 proteins in 228 species: Archae - 19; Bacteria - 379; Metazoa - 79; Fungi - 4; Plants - 29; Viruses - 0; Other Eukaryotes - 133 (source: NCBI BLink). |
| GmaAffx.91156.1.S1_ | 0.05603 | 2.2972 | Glyma14g38890.1 | AT3G46940.1 | deoxyuridine 5'-triphosphate nucleotidohydrolase family; FUNCTIONS IN: hydrolase activity, dUTP diphosphatase activity; INVOLVED IN: 2'-deoxyribonucleotide metabolic process, dUTP metabolic process; LOCATED IN: cellular_component unknown; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: DeoxyUTP pyrophosphatase subfamily 1 (InterPro:IPR008181), DeoxyUTP pyrophosphatase (InterPro:IPR008180); Has 5469 Blast hits to 5455 proteins in 1432 species: Archae - 13; Bacteria - 2270; Metazoa - 201; Fungi - 106; Plants - 30; Viruses - 721; Other Eukaryotes - 2128 (source: NCBI BLink). |
| Gma.12605.1.A1_at | 0.0472 | 2.2958 | Glyma10g32440.1 | AT3G23280.2 | zinc finger (C3HC4-type RING finger) family protein / ankyrin repeat family protein; FUNCTIONS IN: protein binding, zinc ion binding; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Zinc finger, RING-type (InterPro:IPR001841), Zinc finger, RING/FYVE/PHD-type (InterPro:IPR013083), Ankyrin (InterPro:IPR002110); BEST Arabidopsis thaliana protein match is: zinc finger (C3HC4-type RING finger) family protein / ankyrin repeat family protein (TAIR:AT4G14365.1); Has 22221 Blast hits to 12172 proteins in 569 species: Archae - 39; Bacteria - 1262; Metazoa - 12213; Fungi - 1009; Plants - 663; Viruses - 384; Other Eukaryotes - 6651 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.2560.3.S1_at | 0.07776 | 2.2555 | Glyma07g36440.1 | AT3G06500.1 | beta-fructofuranosidase, putative / invertase, putative / saccharase, putative / beta-fructosidase, putative; FUNCTIONS IN: catalytic activity, beta-fructofuranosidase activity; INVOLVED IN: biological_process unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Plant neutral invertase (InterPro:IPR006937), Six-hairpin glycosidase-like (InterPro:IPR008928); BEST Arabidopsis thaliana protein match is: beta-fructofuranosidase, putative / invertase, putative / saccharase, putative / beta-fructosidase, putative (TAIR:AT1G56560.1); Has 513 Blast hits to 512 proteins in 77 species: Archae - 0; Bacteria - 104; Metazoa - 0; Fungi - 0; Plants - 173; Viruses - 0; Other Eukaryotes - 236 (source: NCBI BLink). |
| GmaAffx.78365.1.S1_ | 0.04141 | 2.2032 | Glyma07g15930.2 | AT5G02100.1 | Encodes a protein that binds to beta-sitosterol and localizes to the ER. The WFDE motif in ORP3a appears to be important for a direct interaction with PVA12 [Plant VAMP-Associated protein 12]. Mutation of this motif causes ORP3a to relocalize to the Golgi and cytosol. The interaction between PVA12 and ORP3a does not appear to be sterol-dependent. |
| GmaAffx.64509.1.S1_ | 0.08725 | 2.1979 | Glyma15g13620.1 | AT5G04885.1 | glycosyl hydrolase family 3 protein; FUNCTIONS IN: hydrolase activity, hydrolyzing O-glycosyl compounds; INVOLVED IN: carbohydrate metabolic process; LOCATED IN: plasma membrane, anchored to membrane; EXPRESSED IN: 7 plant structures; CONTAINS InterPro DOMAIN/s: Glycoside hydrolase, family 3, N-terminal (InterPro:IPR001764), Glycoside hydrolase, family 3, C-terminal (InterPro:IPR002772), Glycoside hydrolase, catalytic core (InterPro:IPR017853); BEST Arabidopsis thaliana protein match is: glycosyl hydrolase family 3 protein (TAIR:AT5G20950.2); Has 5829 Blast hits to 5471 proteins in 844 species: Archae - 20; Bacteria - 2907; Metazoa - 6; Fungi - 873; Plants - 279; Viruses - 0; Other Eukaryotes - 1744 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.4016.1.S1_at | 0.04522 | 2.1974 | Glyma18g38350.1 | AT4G30480.2 | tetratricopeptide repeat (TPR)-containing protein; FUNCTIONS IN: binding; LOCATED IN: cellular_component unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Tetratricopeptide TPR-1 (InterPro:IPR001440), Tetratricopeptide-like helical (InterPro:IPR011990), Tetratricopeptide region (InterPro:IPR013026); BEST Arabidopsis thaliana protein match is: TTL4 (Tetratricopetide-repeat Thioredoxin-Like 4); binding (TAIR:AT3G58620.1); Has 10494 Blast hits to 7864 proteins in 509 species: Archae - 360; Bacteria - 1926; Metazoa - 3500; Fungi - 950; Plants - 1004; Viruses - 24; Other Eukaryotes - 2730 (source: NCBI BLink). |
| Gma.7061.1.S1_at | 0.03947 | 2.1962 | Glyma11g17520.1 | AT2G02580.1 | member of CYP71B |
| GmaAffx.399.1.S1_at | 0.08139 | 2.1695 | Glyma15g04890.1 | AT3G45740.1 | hydrolase family protein / HAD-superfamily protein; FUNCTIONS IN: hydrolase activity; INVOLVED IN: metabolic process; LOCATED IN: mitochondrion; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: HAD-superfamily hydrolase, subfamily IIA, CECR5 (InterPro:IPR006353), HAD-superfamily hydrolase, subfamily IIA (InterPro:IPR006357); Has 367 Blast hits to 355 proteins in 105 species: Archae - 5; Bacteria - 2; Metazoa - 102; Fungi - 197; Plants - 14; Viruses - 0; Other Eukaryotes - 47 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.2588.1.S1_at | 0.06392 | 2.0944 | Glyma18g52070.1 | AT1G27980.1 | DPL1; FUNCTIONS IN: pyridoxal phosphate binding, carboxy-lyase activity, catalytic activity; INVOLVED IN: amino acid metabolic process, sphingolipid catabolic process; LOCATED IN: endoplasmic reticulum, membrane; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Pyridoxal phosphate-dependent transferase, major region (InterPro:IPR015424), Pyridoxal phosphate-dependent decarboxylase (InterPro:IPR002129), Pyridoxal phosphate-dependent transferase, major region, subdomain 1 (InterPro:IPR015421); BEST Arabidopsis thaliana protein match is: GAD5 (glutamate decarboxylase 5); calmodulin binding (TAIR:AT3G17760.2); Has 3835 Blast hits to 3826 proteins in 964 species: Archae - 162; Bacteria - 2428; Metazoa - 269; Fungi - 314; Plants - 180; Viruses - 3; Other Eukaryotes - 479 (source: NCBI BLink). |
| Gma.7872.2.S1_at | 0.03369 | 2.0472 | Glyma19g01240.1 | AT5G40270.1 | metal-dependent phosphohydrolase HD domain-containing protein; FUNCTIONS IN: catalytic activity; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 12 growth stages; CONTAINS InterPro DOMAIN/s: Metal-dependent phosphohydrolase, HD region, subdomain (InterPro:IPR006674), Metal-dependent phosphohydrolase, HD region (InterPro:IPR003607); BEST Arabidopsis thaliana protein match is: metal-dependent phosphohydrolase HD domain-containing protein (TAIR:AT5G40290.1); Has 1528 Blast hits to 1521 proteins in 476 species: Archae - 168; Bacteria - 743; Metazoa - 146; Fungi - 6; Plants - 20; Viruses - 3; Other Eukaryotes - 442 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.6427.2.S1_a_at | 0.03375 | 2.017 | Glyma04g38000.2 | AT5G61790.1 | calnexin 1 (CNX1); FUNCTIONS IN: unfolded protein binding, calcium ion binding; INVOLVED IN: protein folding; LOCATED IN: in 8 components; EXPRESSED IN: 26 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Calreticulin/calnexin, P (InterPro:IPR009033), Calreticulin/calnexin (InterPro:IPR001580), Calreticulin/calnexin, conserved site (InterPro:IPR018124), Concanavalin A-like lectin/glucanase (InterPro:IPR008985), Concanavalin A-like lectin/glucanase, subgroup (InterPro:IPR013320); BEST Arabidopsis thaliana protein match is: calnexin, putative (TAIR:AT5G07340.1); Has 1344 Blast hits to 1252 proteins in 297 species: Archae - 2; Bacteria - 61; Metazoa - 623; Fungi - 137; Plants - 191; Viruses - 36; Other Eukaryotes - 294 (source: NCBI BLink). |
| Gma.5683.1.S1_s_at | 0.09121 | 2.0055 | Glyma07g06680.1 | AT2G04520.1 | eukaryotic translation initiation factor 1A, putative / eIF-1A, putative / eIF-4C, putative; FUNCTIONS IN: RNA binding, translation initiation factor activity; INVOLVED IN: translational initiation; LOCATED IN: cellular_component unknown; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Nucleic acid-binding, OB-fold-like (InterPro:IPR016027), Nucleic acid-binding, OB-fold (InterPro:IPR012340), S1, IF1 type (InterPro:IPR006196), Translation initiation factor 1A (eIF-1A), conserved site (InterPro:IPR018104), Translation initiation factor 1A (eIF-1A) (InterPro:IPR001253); BEST Arabidopsis thaliana protein match is: eukaryotic translation initiation factor 1A, putative / eIF-1A, putative / eIF-4C, putative (TAIR:AT5G35680.2); Has 760 Blast hits to 760 proteins in 211 species: Archae - 192; Bacteria - 0; Metazoa - 195; Fungi - 64; Plants - 47; Viruses - 0; Other Eukaryotes - 262 (source: NCBI BLink). |
| Gma.7719.1.A1_at | 0.03775 | 1.9564 | Glyma20g18880.1 | AT1G10840.1 | Encodes eukaryotic initiation factor 3H1 subunit (TIF3H1). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.10547.1.S1_at | 0.05378 | 1.9527 | Glyma03g41850.1 | AT2G46500.2 | Phosphoinositide kinase which undergo autophosphorylation and phosphorylate serine/threonine residues of protein substrates. Contains phosphoinositide 3/4-kinase and ubiquitin-like domains. Phosphorylates PUFD1 and RPN10 in vitro. |
| Gma.2666.1.S1_at | 0.04317 | 1.9481 | Glyma14g39930.1 | AT1G76130.1 | alpha-amylase, putative / 1,4-alpha-D-glucan glucanohydrolase, putative, strong similarity to alpha-amylase GI:7532799 from (Malus x domestica);contains Pfam profile PF00128: Alpha amylase, catalytic domain. Predicted to be secreted based on SignalP analysis. |
| GmaAffx.63114.1.A1 | 0.04942 | 1.9408 | Glyma08g19140.1 | AT1G78160.1 | Encodes a member of the Arabidopsis Pumilio (APUM) proteins containing PUF domain (eight repeats of approximately 36 amino acids each). PUF proteins regulate both mRNA stability and translation through sequence-specific binding to the 3' UTR of target mRNA transcripts. |
| GmaAffx.6037.1.A1_ | 0.07475 | 1.912 | Glyma02g07840.1 | AT1G64760.2 | glycosyl hydrolase family 17 protein; FUNCTIONS IN: cation binding, hydrolase activity, hydrolyzing O-glycosyl compounds, catalytic activity; INVOLVED IN: carbohydrate metabolic process; LOCATED IN: anchored to plasma membrane, plasma membrane, anchored to membrane; EXPRESSED IN: guard cell, cultured cell; CONTAINS InterPro DOMAIN/s: X8 (InterPro:IPR012946), Glycoside hydrolase, catalytic core (InterPro:IPR017853), Glycoside hydrolase, family 17 (InterPro:IPR000490), Glycoside hydrolase, subgroup, catalytic core (InterPro:IPR013781); BEST Arabidopsis thaliana protein match is: glycosyl hydrolase family 17 protein (TAIR:AT2G19440.1); Has 1676 Blast hits to 1621 proteins in 125 species: Archae - 0; Bacteria - 4; Metazoa - 0; Fungi - 26; Plants - 1637; Viruses - 0; Other Eukaryotes - 9 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.6427.1.S1_x_at | 0.05809 | 1.9091 | Glyma06g17060.2 | AT5G61790.1 | calnexin 1 (CNX1); FUNCTIONS IN: unfolded protein binding, calcium ion binding; INVOLVED IN: protein folding; LOCATED IN: in 8 components; EXPRESSED IN: 26 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Calreticulin/calnexin, P (InterPro:IPR009033), Calreticulin/calnexin (InterPro:IPR001580), Calreticulin/calnexin, conserved site (InterPro:IPR018124), Concanavalin A-like lectin/glucanase (InterPro:IPR008985), Concanavalin A-like lectin/glucanase, subgroup (InterPro:IPR013320); BEST Arabidopsis thaliana protein match is: calnexin, putative (TAIR:AT5G07340.1); Has 1344 Blast hits to 1252 proteins in 297 species: Archae - 2; Bacteria - 61; Metazoa - 623; Fungi - 137; Plants - 191; Viruses - 36; Other Eukaryotes - 294 (source: NCBI BLink). |
| Gma.14182.2.A1_at | 0.06847 | 1.9063 | Glyma12g35830.1 | AT1G16900.1 | curculin-like (mannose-binding) lectin family protein, very low similarity to Ser Thr protein kinase GI:2598067 from (Zea mays); contains Pfam lectin (probable mannose binding) domain PF01453 but not the protein kinase domain of the Z. mays protein |
| GmaAffx.42589.1.A1 | 0.07488 | 1.8711 | Glyma09g12030.1 | AT2G12550.1 | ubiquitin-associated (UBA)/TS-N domain-containing protein; INVOLVED IN: protein modification process; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Ubiquitin-associated/translation elongation factor EF1B, N-terminal, eukaryote (InterPro:IPR015940), Ubiquitin-associated/translation elongation factor EF1B, N-terminal (InterPro:IPR000449), Ubiquitin (InterPro:IPR000626), UBA-like (InterPro:IPR009060); Has 274 Blast hits to 274 proteins in 64 species: Archae - 0; Bacteria - 8; Metazoa - 191; Fungi - 2; Plants - 23; Viruses - 0; Other Eukaryotes - 50 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.3585.1.S1_a_at | 0.03473 | 1.8589 | Glyma07g19460.1 | AT2G02090.1 | ETL1; FUNCTIONS IN: helicase activity, DNA binding, ATP binding, nucleic acid binding; LOCATED IN: chloroplast; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: DEAD-like helicase, N-terminal (InterPro:IPR014001), DNA/RNA helicase, C-terminal (InterPro:IPR001650), Helicase, superfamily 1 and 2, ATP-binding (InterPro:IPR014021), SNF2-related (InterPro:IPR000330); BEST Arabidopsis thaliana protein match is: CHR11 (CHROMATIN-REMODELING PROTEIN 11); ATP binding / DNA binding / DNA-dependent ATPase/ helicase/ hydrolase, acting on acid anhydrides, in phosphorus-containing anhydrides / nucleic acid binding / nucleosome binding (TAIR:AT3G06400.1); Has 13578 Blast hits to 11155 proteins in 1163 species: Archae - 82; Bacteria - 3430; Metazoa - 3239; Fungi - 2640; Plants - 921; Viruses - 195; Other Eukaryotes - 3071 (source: NCBI BLink). |
| GmaAffx.10164.1.A1 | 0.04759 | 1.8518 | Glyma14g09680.1 | AT4G24230.6 | acyl-CoA-binding protein ACBP3. Localized extracellularly in transiently expressed tobacco BY-2 cells and onion epidermal cells. Binds arachidonyl-CoA with high affinity. |

| | | | | | |
|---|---|---|---|---|---|
| Gma.17616.1.S1_at | 0.06818 | 1.8516 | Glyma10g35700.1 | AT5G60390.3 | elongation factor 1-alpha / EF-1-alpha; FUNCTIONS IN: calmodulin binding, translation elongation factor activity; INVOLVED IN: translational elongation; LOCATED IN: mitochondrion, nucleus, cytoplasm; EXPRESSED IN: cotyledon, male gametophyte, guard cell, pollen tube, seed; EXPRESSED DURING: L mature pollen stage, M germinated pollen stage; CONTAINS InterPro DOMAIN/s: Translation elongation factor EFTu/EF1A, C-terminal (InterPro:IPR004160), Translation elongation factor EFTu/EF1A, domain 2 (InterPro:IPR004161), Translation elongation factor EF1A/initiation factor IF2gamma, C-terminal (InterPro:IPR009001), Protein synthesis factor, GTP-binding (InterPro:IPR000795), Translation elongation and initiation factors/Ribosomal, beta-barrel (InterPro:IPR009000), Translation elongation factor EF1A, eukaryotic and archaeal (InterPro:IPR004539); BEST Arabidopsis thaliana protein match is: elongation factor 1-alpha / EF-1-alpha (TAIR:AT1G07940.2); Has 56847 Blast hits to 56774 proteins in 13741 species: Archae - 656; Bacteria - 18823; Metazoa - 15489; Fungi - 8699; Plants - 1228; Viruses - 3; Other Eukaryotes - 11949 (source: NCBI BLink |
| GmaAffx.23322.1.S1_ | 0.05794 | 1.8268 | Ambiguous Hit | AT2G26310.1 | chalcone isomerase; FUNCTIONS IN: chalcone isomerase activity; INVOLVED IN: flavonoid biosynthetic process; LOCATED IN: cellular_component unknown; CONTAINS InterPro DOMAIN/s: Chalcone isomerase, subgroup (InterPro:IPR003466), Chalcone isomerase (InterPro:IPR016087); BEST Arabidopsis thaliana protein match is: chalcone isomerase (TAIR:AT3G63170.1); Has 65 Blast hits to 65 proteins in 15 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 4; Plants - 59; Viruses - 0; Other Eukaryotes - 2 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.13471.1.A1_at | 0.05376 | 1.8047 | Glyma08g12020.1 | AT3G13560.2 | glycosyl hydrolase family 17 protein; FUNCTIONS IN: cation binding, hydrolase activity, hydrolyzing O-glycosyl compounds, catalytic activity; INVOLVED IN: carbohydrate metabolic process; LOCATED IN: anchored to plasma membrane, plasma membrane, anchored to membrane; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: X8 (InterPro:IPR012946), Glycoside hydrolase, catalytic core (InterPro:IPR017853), Glycoside hydrolase, family 17 (InterPro:IPR000490), Glycoside hydrolase, subgroup, catalytic core (InterPro:IPR013781); BEST Arabidopsis thaliana protein match is: glycosyl hydrolase family 17 protein / beta-1,3-glucanase, putative (TAIR:AT2G01630.1); Has 1780 Blast hits to 1728 proteins in 123 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 32; Plants - 1744; Viruses - 0; Other Eukaryotes - 4 (source: NCBI BLink). |
| GmaAffx.69277.1.S1_ | 0.06499 | 1.7835 | Glyma20g05620.1 | AT2G02970.1 | nucleoside phosphatase family protein / GDA1/CD39 family protein; FUNCTIONS IN: hydrolase activity; INVOLVED IN: biological_process unknown; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Nucleoside phosphatase GDA1/CD39 (InterPro:IPR000407); BEST Arabidopsis thaliana protein match is: nucleoside phosphatase family protein / GDA1/CD39 family protein (TAIR:AT1G14230.1); Has 1085 Blast hits to 1080 proteins in 167 species: Archae - 0; Bacteria - 20; Metazoa - 529; Fungi - 212; Plants - 199; Viruses - 0; Other Eukaryotes - 125 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.89435.1.A1 | 0.03154 | 1.765 | Glyma03g38190.2 | AT4G01850.2 | S-ADENOSYLMETHIONINE SYNTHETASE 2 (SAM-2); FUNCTIONS IN: methionine adenosyltransferase activity; INVOLVED IN: one-carbon compound metabolic process, S-adenosylmethionine biosynthetic process; EXPRESSED IN: 27 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: S-adenosylmethionine synthetase (InterPro:IPR002133); BEST Arabidopsis thaliana protein match is: SAM1 (S-ADENOSYLMETHIONINE SYNTHETASE 1); methionine adenosyltransferase (TAIR:AT1G02500.2); Has 8027 Blast hits to 8026 proteins in 1640 species: Archae - 6; Bacteria - 2999; Metazoa - 309; Fungi - 111; Plants - 502; Viruses - 0; Other Eukaryotes - 4100 (source: NCBI BLink). |
| Gma.16400.1.S1_a_at | 0.05941 | 1.7541 | No Soybean Match Identified, E<10E-30, Perecent ID>95% | AT3G07160.1 | Encodes GSL10, a member of the Glucan Synthase-Like (GSL) family believed to be involved in the synthesis of the cell wall component callose. GSL10 is required for male gametophyte development and plant growth. Has a role in entry of microspores into mitosis. GSL10 mutation leads to perturbation of microspore division symmetry, irregular callose deposition and failure of generative cell engulfment by the vegetative cell cytoplasm. Also refer to GSL8 (At2g36850). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.54673.1.S1_ | 0.02679 | 1.7491 | Glyma06g20030.1 | AT5G62620.1 | galactosyltransferase family protein; FUNCTIONS IN: transferase activity, transferring hexosyl groups, sugar binding, galactosyltransferase activity; INVOLVED IN: protein amino acid glycosylation; LOCATED IN: chloroplast; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Galectin, carbohydrate recognition domain (InterPro:IPR001079), Glycosyl transferase, family 31 (InterPro:IPR002659), Concanavalin A-like lectin/glucanase (InterPro:IPR008985), Concanavalin A-like lectin/glucanase, subgroup (InterPro:IPR013320); BEST Arabidopsis thaliana protein match is: galactosyltransferase family protein (TAIR:AT1G74800.1); Has 1757 Blast hits to 1750 proteins in 93 species: Archae - 0; Bacteria - 4; Metazoa - 1402; Fungi - 6; Plants - 315; Viruses - 0; Other Eukaryotes - 30 (source: NCBI BLink). |
| Gma.1928.1.S1_at | 0.05226 | 1.7351 | Glyma03g24600.3 | AT5G26667.3 | encodes a uridine 5'-monophosphate (UMP)/cytidine 5'-monophosphate (CMP) kinase. |
| GmaAffx.28750.1.S1_ | 0.06412 | 1.7347 | Glyma13g22640.2 | AT2G26600.2 | glycosyl hydrolase family 17 protein; FUNCTIONS IN: cation binding, hydrolase activity, hydrolyzing O-glycosyl compounds, catalytic activity; INVOLVED IN: carbohydrate metabolic process; LOCATED IN: anchored to membrane; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Glycoside hydrolase, catalytic core (InterPro:IPR017853), Glycoside hydrolase, family 17 (InterPro:IPR000490), Glycoside hydrolase, subgroup, catalytic core (InterPro:IPR013781); BEST Arabidopsis thaliana protein match is: glycosyl hydrolase family 17 protein (TAIR:AT3G15800.1); Has 1369 Blast hits to 1360 proteins in 111 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 3; Plants - 1362; Viruses - 0; Other Eukaryotes - 4 (source: NCBI BLink). |
| GmaAffx.23763.1.S1_ | 0.07148 | 1.7213 | Glyma05g31390.1 | AT5G65110.1 | Encodes an acyl-CoA oxidase presumably involved in long chain fatty acid biosynthesis. |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.21003.1.S1_ | 0.06496 | 1.7206 | Glyma13g33970.2 | AT1G78580.1 | Encodes an enzyme putatively involved in trehalose biosynthesis. The protein has a trehalose synthase (TPS)-like domain but no trehalose phosphatase (TPP)-like domain. ATTPS1 is able to complement yeast tps1 mutants in vivo. The gene product modulates cell growth but not cell differentiation by determining cell wall deposition and cell division. |
| Gma.2560.3.S1_a_at | 0.07219 | 1.7152 | Glyma07g36440.1 | AT3G06500.1 | beta-fructofuranosidase, putative / invertase, putative / saccharase, putative / beta-fructosidase, putative; FUNCTIONS IN: catalytic activity, beta-fructofuranosidase activity; INVOLVED IN: biological_process unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Plant neutral invertase (InterPro:IPR006937), Six-hairpin glycosidase-like (InterPro:IPR008928); BEST Arabidopsis thaliana protein match is: beta-fructofuranosidase, putative / invertase, putative / saccharase, putative / beta-fructosidase, putative (TAIR:AT1G56560.1); Has 513 Blast hits to 512 proteins in 77 species: Archae - 0; Bacteria - 104; Metazoa - 0; Fungi - 0; Plants - 173; Viruses - 0; Other Eukaryotes - 236 (source: NCBI BLink). |
| GmaAffx.92540.1.S1_ | 0.09251 | 1.6959 | Glyma16g08390.2 | AT4G19180.1 | nucleoside phosphatase family protein / GDA1/CD39 family protein; FUNCTIONS IN: hydrolase activity; EXPRESSED IN: cultured cell; CONTAINS InterPro DOMAIN/s: Nucleoside phosphatase GDA1/CD39 (InterPro:IPR000407); BEST Arabidopsis thaliana protein match is: nucleoside phosphatase family protein / GDA1/CD39 family protein (TAIR:AT2G02970.1); Has 1018 Blast hits to 1017 proteins in 158 species: Archae - 0; Bacteria - 20; Metazoa - 512; Fungi - 202; Plants - 187; Viruses - 0; Other Eukaryotes - 97 (source: NCBI BLink). |
| Gma.7192.1.S1_at | 0.06659 | 1.6588 | Glyma16g08960.3 | AT5G65940.1 | hydrolyzes beta-hydroxyisobutyryl-CoA |
| Gma.16749.1.S1_at | 0.04065 | 1.6554 | Glyma10g01960.1 | AT3G62660.1 | Encodes a protein with putative galacturonosyltransferase activity. |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.26289.1.S1_ | 0.06337 | 1.633 | Glyma10g11060.1 | AT5G06570.2 | hydrolase; FUNCTIONS IN: hydrolase activity; INVOLVED IN: metabolic process; LOCATED IN: cellular_component unknown; EXPRESSED IN: 6 plant structures; EXPRESSED DURING: petal differentiation and expansion stage; CONTAINS InterPro DOMAIN/s: Alpha/beta hydrolase fold-3 (InterPro:IPR013094); BEST Arabidopsis thaliana protein match is: AtCXE17 (Arabidopsis thaliana carboxyesterase 17); hydrolase (TAIR:AT5G16080.1); Has 6488 Blast hits to 6477 proteins in 918 species: Archae - 49; Bacteria - 3313; Metazoa - 668; Fungi - 664; Plants - 658; Viruses - 3; Other Eukaryotes - 1133 (source: NCBI BLink). |
| GmaAffx.85865.1.S1_ | 0.06392 | 1.6285 | Glyma19g10160.2 | AT5G47750.1 | D6 PROTEIN KINASE LIKE 2 (D6PKL2); FUNCTIONS IN: kinase activity; INVOLVED IN: protein amino acid phosphorylation; LOCATED IN: nucleolus, nucleus, cytoplasm; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Protein kinase, core (InterPro:IPR000719), Serine/threonine protein kinase (InterPro:IPR002290), Serine/threonine protein kinase-related (InterPro:IPR017442), Protein kinase-like (InterPro:IPR011009), Serine/threonine protein kinase, active site (InterPro:IPR008271); BEST Arabidopsis thaliana protein match is: D6PKL1 (D6 PROTEIN KINASE LIKE 1); kinase (TAIR:AT4G26610.1); Has 86667 Blast hits to 62902 proteins in 1914 species: Archae - 19; Bacteria - 7097; Metazoa - 40115; Fungi - 8950; Plants - 11718; Viruses - 311; Other Eukaryotes - 18457 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.8247.2.S1_a | 0.07755 | 1.6024 | Glyma06g30880.1 | AT5G57870.2 | eukaryotic translation initiation factor 4F, putative / eIF-4F, putative; FUNCTIONS IN: protein binding, RNA binding, binding, translation initiation factor activity; INVOLVED IN: translational initiation, translation, RNA metabolic process; LOCATED IN: nucleus, cytoplasm; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Initiation factor eIF-4 gamma, MA3 (InterPro:IPR003891), Armadillo-type fold (InterPro:IPR016024), MIF4G-like, type 3 (InterPro:IPR003890), MIF4-like, type 1/2/3 (InterPro:IPR016021); BEST Arabidopsis thaliana protein match is: MIF4G domain-containing protein / MA3 domain-containing protein (TAIR:AT2G24050.1); Has 1656 Blast hits to 1406 proteins in 174 species: Archae - 0; Bacteria - 18; Metazoa - 1007; Fungi - 128; Plants - 289; Viruses - 0; Other Eukaryotes - 214 (source: NCBI BLink). |
| Gma.6298.1.S1_at | 0.06786 | 1.596 | No Soybean Match Identified, E<10E-30, Perecent ID>95% | AT3G62290.1 | A member of ARF GTPase family. A thaliana has 21 members of this family, known to be essential for vesicle coating and uncoating and functions in GTP-binding. The gene is shown to play a role in cell division, cell expansion and cellulose production using antisense construct. |
| GmaAffx.13133.1.S1_ | 0.05952 | 1.5855 | Glyma18g22680.1 | AT5G28050.2 | cytidine/deoxycytidylate deaminase family protein; FUNCTIONS IN: hydrolase activity, zinc ion binding, catalytic activity; LOCATED IN: cellular_component unknown; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: APOBEC/CMP deaminase, zinc-binding (InterPro:IPR016192), CMP/dCMP deaminase, zinc-binding (InterPro:IPR002125), Cytidine deaminase-like (InterPro:IPR016193); BEST Arabidopsis thaliana protein match is: cytidine/deoxycytidylate deaminase family protein (TAIR:AT3G05300.1); Has 5055 Blast hits to 5055 proteins in 1279 species: Archae - 43; Bacteria - 2975; Metazoa - 92; Fungi - 125; Plants - 79; Viruses - 0; Other Eukaryotes - 1741 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.13697.1.A1_at | 0.02727 | 1.5729 | Glyma15g13620.1 | AT5G04885.1 | glycosyl hydrolase family 3 protein; FUNCTIONS IN: hydrolase activity, hydrolyzing O-glycosyl compounds; INVOLVED IN: carbohydrate metabolic process; LOCATED IN: plasma membrane, anchored to membrane; EXPRESSED IN: 7 plant structures; CONTAINS InterPro DOMAIN/s: Glycoside hydrolase, family 3, N-terminal (InterPro:IPR001764), Glycoside hydrolase, family 3, C-terminal (InterPro:IPR002772), Glycoside hydrolase, catalytic core (InterPro:IPR017853); BEST Arabidopsis thaliana protein match is: glycosyl hydrolase family 3 protein (TAIR:AT5G20950.2); Has 5829 Blast hits to 5471 proteins in 844 species: Archae - 20; Bacteria - 2907; Metazoa - 6; Fungi - 873; Plants - 279; Viruses - 0; Other Eukaryotes - 1744 (source: NCBI BLink). |
| Gma.6298.1.S1_x_at | 0.06304 | 1.5689 | Glyma08g39360.1 | AT3G62290.1 | A member of ARF GTPase family. A thaliana has 21 members of this family, known to be essential for vesicle coating and uncoating and functions in GTP-binding. The gene is shown to play a role in cell division, cell expansion and cellulose production using antisense construct. |
| GmaAffx.90376.1.A1 | 0.05105 | 1.5658 | Glyma14g01430.2 | AT1G55740.1 | Arabidopsis thaliana seed imbibition 1 (AtSIP1); FUNCTIONS IN: hydrolase activity, hydrolyzing O-glycosyl compounds; LOCATED IN: cellular_component unknown; EXPRESSED IN: male gametophyte, sepal, root, carpel; EXPRESSED DURING: 4 anthesis; CONTAINS InterPro DOMAIN/s: Glycoside hydrolase, catalytic core (InterPro:IPR017853), Raffinose synthase (InterPro:IPR008811); BEST Arabidopsis thaliana protein match is: AtSIP2 (Arabidopsis thaliana seed imbibition 2); hydrolase, hydrolyzing O-glycosyl compounds (TAIR:AT3G57520.1); Has 290 Blast hits to 268 proteins in 85 species: Archae - 9; Bacteria - 46; Metazoa - 0; Fungi - 51; Plants - 178; Viruses - 0; Other Eukaryotes - 6 (source: NCBI BLink). |
| GmaAffx.9545.3.S1_s | 0.05762 | 1.5537 | Glyma09g01980.1 | AT5G47780.1 | Encodes a protein with putative galacturonosyltransferase activity. |

| | | | | | |
|---|---|---|---|---|---|
| Gma.7631.1.A1_at | 0.09251 | 1.551 | Glyma06g09220.3 | AT5G25880.1 | The malic enzyme (EC 1.1.1.40) encoded by the ATNADP-ME3 is presumably cytosolic and restricted in its expression by both developmental and cell-specific signals. |
| GmaAffx.82609.1.S1_ | 0.08312 | 1.5391 | Glyma03g03280.1 | AT1G02020.1 | nitroreductase family protein; FUNCTIONS IN: oxidoreductase activity, acting on NADH or NADPH, nitrogenous group as acceptor, oxidoreductase activity; INVOLVED IN: metabolic process; LOCATED IN: chloroplast; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Nitroreductase (InterPro:IPR000415); Has 465 Blast hits to 402 proteins in 182 species: Archae - 87; Bacteria - 350; Metazoa - 0; Fungi - 0; Plants - 11; Viruses - 0; Other Eukaryotes - 17 (source: NCBI BLink). |
| GmaAffx.91183.1.S1_ | 0.03377 | 1.5234 | Glyma19g06470.4 | AT5G16120.2 | hydrolase, alpha/beta fold family protein; FUNCTIONS IN: hydrolase activity; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Alpha/beta hydrolase fold-1 (InterPro:IPR000073); BEST Arabidopsis thaliana protein match is: hydrolase, alpha/beta fold family protein (TAIR:AT1G77420.1); Has 3068 Blast hits to 3064 proteins in 815 species: Archae - 20; Bacteria - 1988; Metazoa - 126; Fungi - 101; Plants - 252; Viruses - 41; Other Eukaryotes - 540 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.781.1.S1_at | 0.07755 | 1.5142 | Glyma18g04540.1 | AT5G37710.1 | lipase class 3 family protein / calmodulin-binding heat-shock protein, putative; FUNCTIONS IN: triacylglycerol lipase activity, calmodulin binding; INVOLVED IN: lipid catabolic process, lipid metabolic process; LOCATED IN: plasma membrane; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Lipase, class 3 (InterPro:IPR002921), Mono- and diacylglycerol lipase, N terminal (InterPro:IPR005592); BEST Arabidopsis thaliana protein match is: lipase class 3 family protein / calmodulin-binding heat-shock protein, putative (TAIR:AT3G49050.1); Has 337 Blast hits to 336 proteins in 85 species: Archae - 0; Bacteria - 0; Metazoa - 142; Fungi - 21; Plants - 123; Viruses - 0; Other Eukaryotes - 51 (source: NCBI BLink). |
| Gma.6298.1.S1_s_at | 0.03473 | 1.5117 | No Soybean Match Identified, E<10E-30, Perecent ID>95% | AT3G62290.1 | A member of ARF GTPase family. A thaliana has 21 members of this family, known to be essential for vesicle coating and uncoating and functions in GTP-binding. The gene is shown to play a role in cell division, cell expansion and cellulose production using antisense construct. |
| GmaAffx.41283.1.S1_ | 0.05588 | 1.5039 | Glyma09g05000.1 | AT3G06930.1 | Encodes an type I protein arginine methyltransferase. PRMT4b can catalyze the asymmetric dimethylation of arginines 2,17, and 26 on histone 3 and can also methylate myelin basic protein in vitro. Double mutants lacking PRMT4a and 4b have reduced levels of histone 3 methylated at R17. These double mutants flower late due to defects in the autonomous pathway and they have elevated levels of FLC transcripts. |
| Gma.2110.1.S1_x_at | 0.05811 | 1.4792 | Glyma13g19500.4 | AT3G53620.1 | Encodes a soluble protein with inorganic pyrophosphatase activity that is highly specific for Mg-inorganic pyrophosphate. |
| GmaAffx.49163.1.S1_ | 0.0819 | 1.4784 | No Soybean Match Identified, E<10E-30, Perecent ID>95% | AT3G05590.1 | Encodes cytoplasmic ribosomal protein L18. |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.4920.1.S1_a | 0.08324 | 1.4729 | Glyma17g32180.1 | AT4G32120.1 | galactosyltransferase family protein; FUNCTIONS IN: transferase activity, transferring hexosyl groups, galactosyltransferase activity; INVOLVED IN: protein amino acid glycosylation; LOCATED IN: membrane; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Glycosyl transferase, family 31 (InterPro:IPR002659); BEST Arabidopsis thaliana protein match is: transferase, transferring glycosyl groups / transferase, transferring hexosyl groups (TAIR:AT2G25300.1); Has 371 Blast hits to 370 proteins in 60 species: Archae - 0; Bacteria - 0; Metazoa - 147; Fungi - 0; Plants - 220; Viruses - 0; Other Eukaryotes - 4 (source: NCBI BLink). |
| Gma.7240.1.A1_at | 0.08453 | 1.4563 | Glyma14g24770.1 | AT3G49400.1 | transducin family protein / WD-40 repeat family protein; FUNCTIONS IN: nucleotide binding; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: sperm cell, embryo, flower, cultured cell; EXPRESSED DURING: petal differentiation and expansion stage, D bilateral stage; CONTAINS InterPro DOMAIN/s: WD40 repeat-like (InterPro:IPR011046), WD40 repeat, region (InterPro:IPR017986), WD40 repeat (InterPro:IPR001680), WD40/YVTN repeat-like (InterPro:IPR015943); Has 724 Blast hits to 645 proteins in 140 species: Archae - 0; Bacteria - 177; Metazoa - 195; Fungi - 136; Plants - 78; Viruses - 0; Other Eukaryotes - 138 (source: NCBI BLink). |
| Gma.12463.1.S1_at | 0.05603 | 1.4553 | Glyma19g18590.1 | AT3G27160.1 | GHS1 encodes plastid ribosomal protein S21 |

| | | | | | |
|---|---|---|---|---|---|
| Gma.4632.1.A1_at | 0.05804 | 1.4487 | Glyma18g01740.1 | AT1G12990.1 | glycosyl transferase family 17 protein; FUNCTIONS IN: transferase activity, transferring glycosyl groups, acetylglucosaminyltransferase activity; INVOLVED IN: protein amino acid N-linked glycosylation; LOCATED IN: membrane; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Glycosyl transferase, family 17 (InterPro:IPR006813); BEST Arabidopsis thaliana protein match is: glycosyl transferase family 17 protein (TAIR:AT1G67880.1); Has 972 Blast hits to 971 proteins in 58 species: Archae - 0; Bacteria - 24; Metazoa - 46; Fungi - 23; Plants - 68; Viruses - 4; Other Eukaryotes - 807 (source: NCBI BLink). |
| Gma.7092.1.S1_at | 0.05282 | 1.4452 | Glyma02g46330.1 | AT5G55180.1 | glycosyl hydrolase family 17 protein; FUNCTIONS IN: cation binding, hydrolase activity, hydrolyzing O-glycosyl compounds, catalytic activity; INVOLVED IN: carbohydrate metabolic process; LOCATED IN: endomembrane system; EXPRESSED IN: 18 plant structures; EXPRESSED DURING: 11 growth stages; CONTAINS InterPro DOMAIN/s: X8 (InterPro:IPR012946), Glycoside hydrolase, catalytic core (InterPro:IPR017853), Glycoside hydrolase, family 17 (InterPro:IPR000490), Glycoside hydrolase, subgroup, catalytic core (InterPro:IPR013781); BEST Arabidopsis thaliana protein match is: catalytic/ cation binding / hydrolase, hydrolyzing O-glycosyl compounds (TAIR:AT4G26830.1); Has 1770 Blast hits to 1712 proteins in 115 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 8; Plants - 1757; Viruses - 0; Other Eukaryotes - 5 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.38499.1.S1_ | 0.06847 | 1.4294 | Glyma08g10920.1 | AT1G67410.1 | exostosin family protein; FUNCTIONS IN: catalytic activity; INVOLVED IN: biological_process unknown; LOCATED IN: endomembrane system, membrane; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Exostosin-like (InterPro:IPR004263); BEST Arabidopsis thaliana protein match is: exostosin family protein (TAIR:AT1G74680.1); Has 713 Blast hits to 709 proteins in 67 species: Archae - 0; Bacteria - 6; Metazoa - 100; Fungi - 4; Plants - 532; Viruses - 0; Other Eukaryotes - 71 (source: NCBI BLink). |
| Gma.16348.2.A1_at | 0.09452 | 1.4259 | Glyma06g14610.1 | AT1G26910.1 | 60S ribosomal protein L10 (RPL10B); FUNCTIONS IN: structural constituent of ribosome; INVOLVED IN: translation; LOCATED IN: ribosome, membrane, chloroplast envelope; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Ribosomal protein L10e (InterPro:IPR001197), Ribosomal protein L10e/L16 (InterPro:IPR016180), Ribosomal protein L10e, conserved site (InterPro:IPR018255); BEST Arabidopsis thaliana protein match is: SAC52 (SUPPRESSOR OF ACAULIS 52); structural constituent of ribosome (TAIR:AT1G14320.1); Has 1313 Blast hits to 1311 proteins in 457 species: Archae - 225; Bacteria - 9; Metazoa - 446; Fungi - 102; Plants - 110; Viruses - 0; Other Eukaryotes - 421 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.90238.1.S1_ | 0.06291 | 1.4161 | Glyma16g27010.3 | AT5G18110.1 | NOVEL CAP-BINDING PROTEIN (NCBP); FUNCTIONS IN: RNA binding, translation initiation factor activity; INVOLVED IN: translational initiation; LOCATED IN: cytoplasm; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Eukaryotic translation initiation factor 4E (eIF-4E) (InterPro:IPR001040); BEST Arabidopsis thaliana protein match is: EIF4E (EUKARYOTIC TRANSLATION INITATION FACTOR 4E); RNA binding / RNA cap binding / protein binding / translation initiation factor (TAIR:AT4G18040.1); Has 1298 Blast hits to 1298 proteins in 204 species: Archae - 0; Bacteria - 0; Metazoa - 608; Fungi - 226; Plants - 212; Viruses - 0; Other Eukaryotes - 252 (source: NCBI BLink). |
| GmaAffx.6569.1.A1_ | 0.08446 | 1.4094 | Glyma20g02070.2 | AT2G01464.1 | |
| GmaAffx.8655.1.S1_a | 0.05952 | 1.3999 | Glyma19g22760.1 | AT3G27320.1 | hydrolase; FUNCTIONS IN: hydrolase activity; INVOLVED IN: metabolic process; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Alpha/beta hydrolase fold-3 (InterPro:IPR013094); BEST Arabidopsis thaliana protein match is: AtCXE16 (Arabidopsis thaliana carboxyesterase 16); hydrolase (TAIR:AT5G14310.1); Has 4326 Blast hits to 4318 proteins in 749 species: Archae - 47; Bacteria - 2238; Metazoa - 264; Fungi - 384; Plants - 631; Viruses - 3; Other Eukaryotes - 759 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.406.3.S1_at | 0.09993 | 1.3922 | No Soybean Match Identified, E<10E-30, Perecent ID>95% | AT4G01850.2 | S-ADENOSYLMETHIONINE SYNTHETASE 2 (SAM-2); FUNCTIONS IN: methionine adenosyltransferase activity; INVOLVED IN: one-carbon compound metabolic process, S-adenosylmethionine biosynthetic process; EXPRESSED IN: 27 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: S-adenosylmethionine synthetase (InterPro:IPR002133); BEST Arabidopsis thaliana protein match is: SAM1 (S-ADENOSYLMETHIONINE SYNTHETASE 1); methionine adenosyltransferase (TAIR:AT1G02500.2); Has 8027 Blast hits to 8026 proteins in 1640 species: Archae - 6; Bacteria - 2999; Metazoa - 309; Fungi - 111; Plants - 502; Viruses - 0; Other Eukaryotes - 4100 (source: NCBI BLink). |
| GmaAffx.84363.1.S1_ | 0.06986 | 1.3919 | Glyma09g24430.1 | AT4G25720.1 | This locus encodes a protein with similarity to gamma-glutamylcyclotransferase that may be involved in catalyzing the formation of pyroglutate residue on proteins that have been post-translationally processed to reveal a glutamine at their N-terminus. Enzymatic assays to test the function of this protein were performed using a truncated form of the protein lacking a signal peptide that is most similar to the AT4G25720.1 protein model. |
| Gma.18016.1.S1_at | 0.08586 | 1.388 | Glyma05g08040.2 | AT2G26890.1 | GRV2 has sequence similarity to the C. elegans protein RME-8 which is involved in endocytosis. grv2 mutants result in a reduction in gravitropic response in hypocotyls and shoots but do not affect root gravitropism. The mutants are defective in amyloplast sedimentation. |
| Gma.14330.2.A1_at | 0.09282 | 1.3867 | Glyma20g28880.1 | AT1G04200.1 | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: N-terminal protein myristoylation; LOCATED IN: cellular_component unknown; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; Has 296 Blast hits to 293 proteins in 83 species: Archae - 0; Bacteria - 0; Metazoa - 238; Fungi - 12; Plants - 16; Viruses - 0; Other Eukaryotes - 30 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.2560.1.S1_at | 0.09622 | 1.3864 | No Soybean Match Identified, E<10E-30, Perecent ID>95% | AT1G56560.1 | beta-fructofuranosidase, putative / invertase, putative / saccharase, putative / beta-fructosidase, putative; FUNCTIONS IN: catalytic activity, beta-fructofuranosidase activity; INVOLVED IN: biological_process unknown; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Plant neutral invertase (InterPro:IPR006937), Six-hairpin glycosidase-like (InterPro:IPR008928); BEST Arabidopsis thaliana protein match is: beta-fructofuranosidase/ catalytic (TAIR:AT3G05820.1); Has 514 Blast hits to 513 proteins in 76 species: Archae - 0; Bacteria - 101; Metazoa - 0; Fungi - 0; Plants - 173; Viruses - 0; Other Eukaryotes - 240 (source: NCBI BLink). |
| Gma.16411.1.S1_at | 0.04164 | 1.3445 | No Soybean Match Identified, E<10E-30, Perecent ID>95% | AT3G61130.1 | Encodes a protein with putative galacturonosyltransferase activity. |
| Gma.17390.1.S1_at | 0.08292 | 1.3162 | Glyma06g46490.1 | AT5G21990.1 | tetratricopeptide repeat (TPR)-containing protein; FUNCTIONS IN: binding; INVOLVED IN: protein folding; LOCATED IN: cellular_component unknown; EXPRESSED IN: cultured cell; CONTAINS InterPro DOMAIN/s: Tetratricopeptide TPR-1 (InterPro:IPR001440), Tetratricopeptide-like helical (InterPro:IPR011990), Tetratricopeptide region (InterPro:IPR013026), Peptidyl-prolyl cis-trans isomerase, FKBP-type (InterPro:IPR001179); BEST Arabidopsis thaliana protein match is: peptidyl-prolyl cis-trans isomerase, putative / FK506-binding protein, putative (TAIR:AT5G48570.1); Has 2799 Blast hits to 2611 proteins in 224 species: Archae - 12; Bacteria - 49; Metazoa - 1583; Fungi - 260; Plants - 352; Viruses - 0; Other Eukaryotes - 543 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.12187.1.A1_s_at | 0.06848 | 1.3134 | Glyma11g33670.1 | AT5G37720.2 | RNA and export factor-binding protein, putative; FUNCTIONS IN: nucleotide binding, nucleic acid binding; INVOLVED IN: biological_process unknown; CONTAINS InterPro DOMAIN/s: RNA recognition motif, RNP-1 (InterPro:IPR000504), Nucleotide-binding, alpha-beta plait (InterPro:IPR012677); BEST Arabidopsis thaliana protein match is: RNA and export factor-binding protein, putative (TAIR:AT1G66260.1); Has 4756 Blast hits to 4444 proteins in 358 species: Archae - 2; Bacteria - 230; Metazoa - 2500; Fungi - 868; Plants - 728; Viruses - 9; Other Eukaryotes - 419 (source: NCBI BLink). |
| Gma.748.1.A1_at | 0.07755 | 1.3037 | Glyma19g06900.1 | AT1G74910.2 | ADP-glucose pyrophosphorylase family protein; FUNCTIONS IN: transferase activity, nucleotidyltransferase activity; INVOLVED IN: biosynthetic process; LOCATED IN: peroxisome; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Trimeric LpxA-like (InterPro:IPR011004), Bacterial transferase hexapeptide repeat (InterPro:IPR001451), Nucleotidyl transferase (InterPro:IPR005835); BEST Arabidopsis thaliana protein match is: ADP-glucose pyrophosphorylase family protein (TAIR:AT2G04650.1); Has 5617 Blast hits to 5611 proteins in 1142 species: Archae - 383; Bacteria - 3084; Metazoa - 330; Fungi - 199; Plants - 216; Viruses - 0; Other Eukaryotes - 1405 (source: NCBI BLink). |
| GmaAffx.43132.1.A1 | 0.09829 | 1.3022 | Glyma07g33530.1 | AT3G19260.1 | LAG1 homolog . Loss of function mutant is sensitive to AAL-toxin. LOH2 is presumed to function in sphingolipid metabolism. |
| Gma.14342.1.S1_at | 0.09454 | 1.2111 | Glyma11g08920.1 | AT4G35260.1 | NAD+ dependent isocitrate dehydrogenase subunit 1 (IDH1) |

| | | | | | |
|---|---|---|---|---|---|
| Gma.10759.1.S1_at | 0.08784 | 1.203 | Glyma01g43470.3 | AT4G23850.1 | long-chain-fatty-acid--CoA ligase / long-chain acyl-CoA synthetase; FUNCTIONS IN: catalytic activity; INVOLVED IN: fatty acid biosynthetic process; LOCATED IN: plasma membrane; EXPRESSED IN: 26 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: AMP-dependent synthetase and ligase (InterPro:IPR000873); BEST Arabidopsis thaliana protein match is: long-chain-fatty-acid--CoA ligase, putative / long-chain acyl-CoA synthetase, putative (TAIR:AT4G11030.1); Has 34929 Blast hits to 33173 proteins in 1967 species: Archae - 465; Bacteria - 18245; Metazoa - 1999; Fungi - 1232; Plants - 1153; Viruses - 1; Other Eukaryotes - 11834 (source: NCBI BLink). |
| GmaAffx.1112.1.S1_a | 0.09706 | 1.1402 | Glyma08g00920.1 | AT5G61790.1 | calnexin 1 (CNX1); FUNCTIONS IN: unfolded protein binding, calcium ion binding; INVOLVED IN: protein folding; LOCATED IN: in 8 components; EXPRESSED IN: 26 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Calreticulin/calnexin, P (InterPro:IPR009033), Calreticulin/calnexin (InterPro:IPR001580), Calreticulin/calnexin, conserved site (InterPro:IPR018124), Concanavalin A-like lectin/glucanase (InterPro:IPR008985), Concanavalin A-like lectin/glucanase, subgroup (InterPro:IPR013320); BEST Arabidopsis thaliana protein match is: calnexin, putative (TAIR:AT5G07340.1); Has 1344 Blast hits to 1252 proteins in 297 species: Archae - 2; Bacteria - 61; Metazoa - 623; Fungi - 137; Plants - 191; Viruses - 36; Other Eukaryotes - 294 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.87100.1.S1_ | 0.06613 | -8.5106 | Glyma12g02330.1 | AT4G10320.1 | isoleucyl-tRNA synthetase, putative / isoleucine--tRNA ligase, putative; FUNCTIONS IN: isoleucine-tRNA ligase activity, nucleotide binding, aminoacyl-tRNA ligase activity, ATP binding; INVOLVED IN: response to cadmium ion, tRNA aminoacylation for protein translation; LOCATED IN: cytosol; EXPRESSED IN: male gametophyte, guard cell, epidermis, cultured cell, pollen tube; EXPRESSED DURING: L mature pollen stage, M germinated pollen stage; CONTAINS InterPro DOMAIN/s: Aminoacyl-tRNA synthetase, class I, conserved site (InterPro:IPR001412), Isoleucyl-tRNA synthetase (InterPro:IPR018353), Isoleucyl-tRNA synthetase, class Ia (InterPro:IPR002301), Aminoacyl-tRNA synthetase, class 1a, anticodon-binding (InterPro:IPR009080), Isoleucyl-tRNA synthetase, class Ia, N-terminal (InterPro:IPR015905), Rossmann-like alpha/beta/alpha sandwich fold (InterPro:IPR014729), Valyl/Leucyl/Isoleucyl-tRNA synthetase, class Ia, editing (InterPro:IPR009008), Valyl/Leucyl/Isoleucyl-tRNA synthetase, class I, anticodon-binding (InterPro:IPR013155), Aminoacyl-tRNA synthetase, class Ia (InterPro:IPR002300); BEST Arabidops |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.66987.1.S1_ | 0.07755 | -7.4123 | Glyma19g07410.2 | AT5G40610.1 | glycerol-3-phosphate dehydrogenase (NAD+) / GPDH; FUNCTIONS IN: in 9 functions; INVOLVED IN: oxidation reduction, glycerol-3-phosphate metabolic process, glycerol-3-phosphate catabolic process, carbohydrate metabolic process, metabolic process; LOCATED IN: chloroplast, glycerol-3-phosphate dehydrogenase complex, cytoplasm; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: 6-phosphogluconate dehydrogenase, C-terminal-like (InterPro:IPR008927), Dehydrogenase, multihelical (InterPro:IPR013328), NAD-dependent glycerol-3-phosphate dehydrogenase, N-terminal (InterPro:IPR011128), NAD-dependent glycerol-3-phosphate dehydrogenase (InterPro:IPR006168), NAD-dependent glycerol-3 phosphate dehydrogenase, eukaryotic (InterPro:IPR017751), NAD(P)-binding (InterPro:IPR016040), NAD-dependent glycerol 3-phosphate dehydrogenase, C-terminal (InterPro:IPR006109); BEST Arabidopsis thaliana protein match is: GLY1; glycerol-3-phosphate dehydrogenase (NAD+) (TAIR:AT2G40690.1); Has 5564 Blast hits to 5552 proteins in 1603 species: Archae - 12; Bacteria - 2810; Metazoa - 565; Fungi - 226; Plants - 97; Viruses - |
| Gma.6550.2.S1_x_at | 0.06733 | -6.9069 | No Soybean Match Identified, E<10E-30, Perecent ID>95% | AT5G54160.1 | A caffeic acid/5-hydroxyferulic acid O-methyltransferase. Interacts with 14-4-3 proteins in yeast 2 hybrid assay. AtOMT1 (At5g54160) encodes a flavonol 3'-O-methyltransferase that is highly active towards quercetin and myricetin. The substrate specificity identifies the enzyme as flavonol 3'-methyltransferase which replaces the former annotation of the gene to encode a caffeic acid/5-hydroxyferulic acid O-methyltransferase |
| Gma.10701.2.S1_at | 0.04466 | -4.8532 | Glyma05g27030.1 | AT3G18270.1 | a cytochrome P450 pseudogene. the second half of the gene overlaps perfectly with the other gene model. |

| | | | | | |
|---|---|---|---|---|---|
| Gma.5491.2.S1_at | 0.09251 | -4.7533 | Glyma09g33820.2 | AT2G23910.1 | cinnamoyl-CoA reductase-related; FUNCTIONS IN: coenzyme binding, binding, cinnamoyl-CoA reductase activity, catalytic activity; INVOLVED IN: lignin biosynthetic process, cellular metabolic process, metabolic process; LOCATED IN: cellular_component unknown; EXPRESSED IN: 19 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: NAD-dependent epimerase/dehydratase (InterPro:IPR001509), NAD(P)-binding (InterPro:IPR016040); BEST Arabidopsis thaliana protein match is: cinnamoyl-CoA reductase-related (TAIR:AT4G30470.1); Has 2416 Blast hits to 2412 proteins in 422 species: Archae - 0; Bacteria - 329; Metazoa - 73; Fungi - 205; Plants - 1191; Viruses - 0; Other Eukaryotes - 618 (source: NCBI BLink). |
| Gma.13775.1.A1_at | 0.06494 | -4.6721 | Glyma20g08830.1 | AT2G19070.1 | encodes a protein whose sequence is similar to anthranilate N-hydroxycinnamoyl/benzoyltransferase from Dianthus caryophyllus (gi:2239091) |
| Gma.3888.2.S1_at | 0.06238 | -4.5102 | Glyma08g21300.1 | AT1G14520.1 | Encodes MIOX1. Belongs to myo-inositol oxygenase gene family. |
| Gma.10591.2.S1_at | 0.08725 | -4.3968 | Glyma11g04530.2 | AT1G32470.1 | glycine cleavage system H protein, mitochondrial, putative; FUNCTIONS IN: glycine dehydrogenase (decarboxylating) activity; INVOLVED IN: glycine catabolic process; LOCATED IN: mitochondrion, glycine cleavage complex, chloroplast; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: 2-oxo acid dehydrogenase, lipoyl-binding site (InterPro:IPR003016), Single hybrid motif (InterPro:IPR011053), Glycine cleavage H-protein (InterPro:IPR002930), Glycine cleavage H-protein, subgroup (InterPro:IPR017453); BEST Arabidopsis thaliana protein match is: GDCH; glycine dehydrogenase (decarboxylating) (TAIR:AT2G35370.1); Has 4756 Blast hits to 4756 proteins in 1228 species: Archae - 82; Bacteria - 2411; Metazoa - 143; Fungi - 84; Plants - 158; Viruses - 0; Other Eukaryotes - 1878 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.8463.1.S1_at | 0.02743 | -3.816 | Glyma18g06950.1 | AT5G05580.1 | Encodes a temperature sensitive plastidic fatty acid desaturase. |
| Gma.7740.1.A1_at | 0.06006 | -3.7493 | Glyma19g40300.1 | AT1G02640.1 | encodes a protein similar to a beta-xylosidase located in the extracellular matrix. This is a member of glycosyl hydrolase family 3 and has six other closely related members. |
| Gma.8015.2.S1_at | 0.05138 | -3.7392 | Glyma11g10980.1 | AT3G19840.1 | Binds the carboxyl-terminal domain (CTD) of the largest subunit of RNA polymerase II and functions as a scaffold for RNA processing machineries. |
| GmaAffx.36820.1.A1 | 0.03272 | -3.5483 | Glyma04g09010.1 | AT2G25790.1 | leucine-rich repeat transmembrane protein kinase, putative; FUNCTIONS IN: protein serine/threonine kinase activity, kinase activity, ATP binding; INVOLVED IN: transmembrane receptor protein tyrosine kinase signaling pathway, protein amino acid phosphorylation; LOCATED IN: plasma membrane; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Protein kinase, core (InterPro:IPR000719), Leucine-rich repeat, N-terminal (InterPro:IPR013210), Leucine-rich repeat (InterPro:IPR001611), Serine/threonine protein kinase-related (InterPro:IPR017442), Protein kinase-like (InterPro:IPR011009); BEST Arabidopsis thaliana protein match is: leucine-rich repeat transmembrane protein kinase, putative (TAIR:AT4G28650.1); Has 99014 Blast hits to 37814 proteins in 1242 species: Archae - 50; Bacteria - 8022; Metazoa - 33960; Fungi - 1313; Plants - 47209; Viruses - 19; Other Eukaryotes - 8441 (source: NCBI BLink). |
| Gma.6550.3.S1_s_at | 0.09858 | -3.2789 | Glyma06g14200.1 | AT5G54160.1 | A caffeic acid/5-hydroxyferulic acid O-methyltransferase. Interacts with 14-4-3 proteins in yeast 2 hybrid assay. AtOMT1 (At5g54160) encodes a flavonol 3'-O-methyltransferase that is highly active towards quercetin and myricetin. The substrate specificity identifies the enzyme as flavonol 3'-methyltransferase which replaces the former annotation of the gene to encode a caffeic acid/5-hydroxyferulic acid O-methyltransferase |

| | | | | | |
|---|---|---|---|---|---|
| Gma.6545.1.S1_at | 0.08935 | -3.1929 | Glyma10g06580.1 | AT5G03300.1 | Encodes adenosine kinase 2 (ADK2), a typical, constitutively expressed housekeeping enzyme. Shows a high sequence identity with ADK1. Involved in salvage synthesis of adenylates and methyl recycling. Enzyme activity is substantially inhibited in roots, siliques and dry seeds by an unknown compound. May contribute to cytokinin interconversion. |
| GmaAffx.8947.1.S1_a | 0.03172 | -3.1448 | Glyma14g05930.1 | AT3G56170.1 | Encodes a calcium-dependent nuclease with similarity to staphylococcal nuclease. |
| GmaAffx.78720.2.S1_ | 0.04512 | -3.1312 | Glyma09g35050.5 | AT4G19420.1 | pectinacetylesterase family protein; FUNCTIONS IN: carboxylesterase activity; INVOLVED IN: biological_process unknown; LOCATED IN: endomembrane system; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Pectinacetylesterase (InterPro:IPR004963); BEST Arabidopsis thaliana protein match is: pectinacetylesterase, putative (TAIR:AT4G19410.1); Has 401 Blast hits to 396 proteins in 86 species: Archae - 2; Bacteria - 40; Metazoa - 114; Fungi - 0; Plants - 161; Viruses - 0; Other Eukaryotes - 84 (source: NCBI BLink). |
| GmaAffx.64904.1.S1_ | 0.07396 | -3.0567 | Glyma17g36650.1 | AT1G19300.1 | The PARVUS/GLZ1 gene encodes a putative family 8 glycosyl transferase that contributes to xylan biosynthesis. Its gene expression shows good co-variance with the IRX3 gene involved in secondary cell wall synthesis. PARVUS/GLZ1 is predicted to have galacturonosyltransferase activity and may be involved in the formation of the complex oligosaccharide sequence present at the reducing end of xylan. PARVUS is expressed in cells undergoing secondary wall thickening, and parvus mutants have thinner cell walls. |
| GmaAffx.31666.1.A1_ | 0.07658 | -2.9802 | Glyma02g47990.1 | AT3G21760.1 | Encodes HYR1, a UDP glycosyltransferase (UGT). HYR1 glucosylates hypostatin, an inhibitor of cell expansion in vivo to form a bioactive glucoside. |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.55753.1.S1_ | 0.08514 | -2.9497 | Glyma14g38860.2 | AT5G17380.1 | pyruvate decarboxylase family protein; FUNCTIONS IN: pyruvate decarboxylase activity, magnesium ion binding, thiamin pyrophosphate binding, transferase activity, catalytic activity; LOCATED IN: cellular_component unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: TPP-binding enzymes, conserved site (InterPro:IPR000399), Thiamine pyrophosphate enzyme, central region (InterPro:IPR012000), Thiamine pyrophosphate enzyme, N-terminal TPP binding region (InterPro:IPR012001), Thiamine pyrophosphate enzyme, C-terminal TPP-binding (InterPro:IPR011766); BEST Arabidopsis thaliana protein match is: CSR1 (CHLORSULFURON/IMIDAZOLINONE RESISTANT 1); acetolactate synthase/ pyruvate decarboxylase (TAIR:AT3G48560.1); Has 19451 Blast hits to 19340 proteins in 1483 species: Archae - 251; Bacteria - 8187; Metazoa - 204; Fungi - 529; Plants - 438; Viruses - 2; Other Eukaryotes - 9840 (source: NCBI BLink). |
| Gma.2075.1.S1_at | 0.08062 | -2.9214 | Glyma07g34690.1 | AT2G35120.1 | glycine cleavage system H protein, mitochondrial, putative; FUNCTIONS IN: glycine dehydrogenase (decarboxylating) activity, ATP binding; INVOLVED IN: glycine catabolic process; LOCATED IN: mitochondrion, glycine cleavage complex; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: 2-oxo acid dehydrogenase, lipoyl-binding site (InterPro:IPR003016), Single hybrid motif (InterPro:IPR011053), Glycine cleavage H-protein (InterPro:IPR002930), Glycine cleavage H-protein, subgroup (InterPro:IPR017453); BEST Arabidopsis thaliana protein match is: glycine cleavage system H protein, mitochondrial, putative (TAIR:AT1G32470.1); Has 4762 Blast hits to 4762 proteins in 1230 species: Archae - 82; Bacteria - 2410; Metazoa - 143; Fungi - 86; Plants - 158; Viruses - 0; Other Eukaryotes - 1883 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.78433.1.S1_ | 0.05811 | -2.837 | Glyma07g33780.2 | AT4G34050.1 | caffeoyl-CoA 3-O-methyltransferase, putative; FUNCTIONS IN: caffeoyl-CoA O-methyltransferase activity; INVOLVED IN: coumarin biosynthetic process, response to cadmium ion; LOCATED IN: cytosol; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: O-methyltransferase, family 3 (InterPro:IPR002935); BEST Arabidopsis thaliana protein match is: caffeoyl-CoA 3-O-methyltransferase, putative (TAIR:AT4G26220.1); Has 3455 Blast hits to 3452 proteins in 740 species: Archae - 19; Bacteria - 1446; Metazoa - 216; Fungi - 70; Plants - 444; Viruses - 0; Other Eukaryotes - 1260 (source: NCBI BLink). |
| Gma.11334.1.S1_a_at | 0.0472 | -2.7654 | No Soybean Match Identified, E<10E-30, Perecent ID>95% | AT4G34050.2 | caffeoyl-CoA 3-O-methyltransferase, putative; FUNCTIONS IN: caffeoyl-CoA O-methyltransferase activity; INVOLVED IN: coumarin biosynthetic process, response to cadmium ion; LOCATED IN: cytosol; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: O-methyltransferase, family 3 (InterPro:IPR002935); BEST Arabidopsis thaliana protein match is: caffeoyl-CoA 3-O-methyltransferase, putative (TAIR:AT4G26220.1); Has 2353 Blast hits to 2341 proteins in 555 species: Archae - 7; Bacteria - 946; Metazoa - 88; Fungi - 25; Plants - 356; Viruses - 0; Other Eukaryotes - 931 (source: NCBI BLink). |
| GmaAffx.3325.1.S1_a | 0.06494 | -2.7363 | Glyma18g05080.1 | AT4G10710.1 | encodes a component of the FAcilitates Chromatin Transcription (FACT) complex, SPT16. |

| | | | | | |
|---|---|---|---|---|---|
| Gma.6062.1.S1_at | 0.04141 | -2.6994 | Glyma03g32130.1 | AT4G14360.2 | dehydration-responsive protein-related; LOCATED IN: Golgi apparatus; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF248, methyltransferase putative (InterPro:IPR004159); BEST Arabidopsis thaliana protein match is: dehydration-responsive protein-related (TAIR:AT3G23300.1); Has 524 Blast hits to 508 proteins in 37 species: Archae - 0; Bacteria - 29; Metazoa - 0; Fungi - 0; Plants - 480; Viruses - 0; Other Eukaryotes - 15 (source: NCBI BLink). |
| GmaAffx.24384.1.A1 | 0.0522 | -2.6721 | Ambiguous Hit | AT5G38530.1 | tryptophan synthase-related; FUNCTIONS IN: pyridoxal phosphate binding, tryptophan synthase activity, catalytic activity; INVOLVED IN: tryptophan biosynthetic process, metabolic process; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Pyridoxal phosphate-dependent enzyme, beta subunit (InterPro:IPR001926), Tryptophan synthase, beta chain-like (InterPro:IPR006316); BEST Arabidopsis thaliana protein match is: TSB2 (TRYPTOPHAN SYNTHASE BETA-SUBUNIT 2); tryptophan synthase (TAIR:AT4G27070.1); Has 7407 Blast hits to 7407 proteins in 1444 species: Archae - 233; Bacteria - 3745; Metazoa - 7; Fungi - 149; Plants - 67; Viruses - 0; Other Eukaryotes - 3206 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.6015.1.A1_at | 0.02679 | -2.6439 | Glyma18g43070.1 | AT3G11210.1 | GDSL-motif lipase/hydrolase family protein; FUNCTIONS IN: hydrolase activity, hydrolase activity, acting on ester bonds, carboxylesterase activity; INVOLVED IN: lipid metabolic process; LOCATED IN: cellular_component unknown; EXPRESSED IN: 20 plant structures; EXPRESSED DURING: 12 growth stages; CONTAINS InterPro DOMAIN/s: Esterase, SGNH hydrolase-type, subgroup (InterPro:IPR013831), Lipase, GDSL (InterPro:IPR001087), Esterase, SGNH hydrolase-type (InterPro:IPR013830); BEST Arabidopsis thaliana protein match is: GDSL-motif lipase/hydrolase family protein (TAIR:AT2G38180.1); Has 410 Blast hits to 409 proteins in 144 species: Archae - 0; Bacteria - 124; Metazoa - 65; Fungi - 98; Plants - 78; Viruses - 0; Other Eukaryotes - 45 (source: NCBI BLink). |
| Gma.340.1.S1_at | 0.03119 | -2.6198 | Glyma15g05720.1 | AT5G64570.1 | Encodes a beta-d-xylosidase that belongs to family 3 of glycoside hydrolases. |
| Gma.5578.1.S1_x_at | 0.04358 | -2.5699 | Glyma19g30600.1 | AT2G40890.1 | encodes coumarate 3-hydroxylase (C3H), a P450-dependent monooxygenase. Involved in lignin biosynthesis and flavonoid biosynthesis. Also affects the biosynthesis of coumarins such as scopoletin and scopolin as a branching-out-pathway from the phenylpropanoid acid level. |
| Gma.5578.3.S1_a_at | 0.03172 | -2.5549 | Glyma03g27740.2 | AT2G40890.1 | encodes coumarate 3-hydroxylase (C3H), a P450-dependent monooxygenase. Involved in lignin biosynthesis and flavonoid biosynthesis. Also affects the biosynthesis of coumarins such as scopoletin and scopolin as a branching-out-pathway from the phenylpropanoid acid level. |

| | | | | | |
|---|---|---|---|---|---|
| Gma.4396.1.S1_at | 0.03172 | -2.5313 | Glyma07g04820.3 | AT3G14200.1 | DNAJ heat shock N-terminal domain-containing protein; FUNCTIONS IN: heat shock protein binding; INVOLVED IN: protein folding; LOCATED IN: cellular_component unknown; EXPRESSED IN: 19 plant structures; EXPRESSED DURING: 11 growth stages; CONTAINS InterPro DOMAIN/s: Molecular chaperone, heat shock protein, Hsp40, DnaJ (InterPro:IPR015609), Heat shock protein DnaJ, N-terminal (InterPro:IPR001623); BEST Arabidopsis thaliana protein match is: heat shock protein binding (TAIR:AT1G72416.3); Has 15240 Blast hits to 15235 proteins in 1880 species: Archae - 101; Bacteria - 5011; Metazoa - 3112; Fungi - 1305; Plants - 1124; Viruses - 13; Other Eukaryotes - 4574 (source: NCBI BLink). |
| Gma.2590.4.S1_a_at | 0.03375 | -2.5145 | Glyma14g01290.1 | AT5G48300.1 | Encodes the small subunit of ADP-glucose pyrophosphorylase. The small subunit is the catalytic isoform responsible for ADP-glucose pyrophosphorylase activity. The presence of the small subunit is required for large subunit stability. Two isoforms of the small subunit (ApS1 and ApS2) have been described. ApS1 is the major small subunit isoform present in all plant tissues tested. |
| GmaAffx.90585.1.S1_ | 0.09324 | -2.4894 | Glyma18g10200.1 | AT5G10770.1 | chloroplast nucleoid DNA-binding protein, putative; FUNCTIONS IN: DNA binding, aspartic-type endopeptidase activity; INVOLVED IN: proteolysis; EXPRESSED IN: 12 plant structures; EXPRESSED DURING: 4 anthesis, C globular stage, petal differentiation and expansion stage, E expanded cotyledon stage; CONTAINS InterPro DOMAIN/s: Peptidase aspartic, catalytic (InterPro:IPR009007), Peptidase A1 (InterPro:IPR001461); BEST Arabidopsis thaliana protein match is: aspartyl protease family protein (TAIR:AT5G10760.1); Has 2665 Blast hits to 2650 proteins in 267 species: Archae - 0; Bacteria - 6; Metazoa - 898; Fungi - 470; Plants - 1168; Viruses - 0; Other Eukaryotes - 123 (source: NCBI BLink). |
| GmaAffx.15813.1.A1 | 0.05719 | -2.4802 | Glyma18g52200.1 | AT2G11890.1 | adenylate cyclase |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.75310.1.S1_ | 0.09212 | -2.4456 | No Soybean Match Identified, E<10E-30, Perecent ID>95% | AT1G75900.1 | family II extracellular lipase 3 (EXL3); FUNCTIONS IN: lipase activity, carboxylesterase activity, acyltransferase activity; INVOLVED IN: sexual reproduction; LOCATED IN: extracellular region; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Lipase, GDSL, active site (InterPro:IPR008265), Lipase, GDSL (InterPro:IPR001087); BEST Arabidopsis thaliana protein match is: family II extracellular lipase 1 (EXL1) (TAIR:AT1G75880.2); Has 1785 Blast hits to 1770 proteins in 149 species: Archae - 0; Bacteria - 196; Metazoa - 1; Fungi - 20; Plants - 1557; Viruses - 0; Other Eukaryotes - 11 (source: NCBI BLink). |
| Gma.2759.1.S1_at | 0.03956 | -2.4377 | Glyma04g37790.1 | AT5G13980.2 | glycosyl hydrolase family 38 protein; FUNCTIONS IN: in 6 functions; INVOLVED IN: mannose metabolic process, carbohydrate metabolic process; LOCATED IN: apoplast, cell wall, plasma membrane, vacuole, plant-type cell wall; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Glycoside hydrolase-type carbohydrate-binding (InterPro:IPR011013), Glycoside hydrolase/deacetylase, beta/alpha-barrel (InterPro:IPR011330), Glycoside hydrolase, family 38, central region (InterPro:IPR015341), Glycoside hydrolase, family 38, core (InterPro:IPR000602), Glycosyl hydrolases 38, C-terminal (InterPro:IPR011682); BEST Arabidopsis thaliana protein match is: glycosyl hydrolase family 38 protein (TAIR:AT3G26720.1); Has 847 Blast hits to 813 proteins in 116 species: Archae - 3; Bacteria - 66; Metazoa - 586; Fungi - 12; Plants - 73; Viruses - 0; Other Eukaryotes - 107 (source: NCBI BLink). |
| Gma.3745.1.S1_at | 0.07712 | -2.3848 | Glyma09g36150.1 | AT5G13700.1 | Encodes a protein with polyamine oxidase activity. The mRNA of this gene is only expressed in very low amounts in the organs where it was detected (light-grown plants). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.89241.1.A1 | 0.05811 | -2.3703 | Glyma17g04340.1 | AT4G01850.2 | S-ADENOSYLMETHIONINE SYNTHETASE 2 (SAM-2); FUNCTIONS IN: methionine adenosyltransferase activity; INVOLVED IN: one-carbon compound metabolic process, S-adenosylmethionine biosynthetic process; EXPRESSED IN: 27 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: S-adenosylmethionine synthetase (InterPro:IPR002133); BEST Arabidopsis thaliana protein match is: SAM1 (S-ADENOSYLMETHIONINE SYNTHETASE 1); methionine adenosyltransferase (TAIR:AT1G02500.2); Has 8027 Blast hits to 8026 proteins in 1640 species: Archae - 6; Bacteria - 2999; Metazoa - 309; Fungi - 111; Plants - 502; Viruses - 0; Other Eukaryotes - 4100 (source: NCBI BLink). |
| Gma.12147.1.S1_at | 0.04816 | -2.3681 | Glyma17g10020.1 | AT5G50090.2 | unknown protein; FUNCTIONS IN: molecular_function unknown; LOCATED IN: cellular_component unknown; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT5G62900.1); Has 99 Blast hits to 99 proteins in 7 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 99; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| Gma.12790.1.S1_at | 0.07755 | -2.3626 | Glyma01g43300.1 | AT1G64185.1 | lactoylglutathione lyase family protein / glyoxalase I family protein; FUNCTIONS IN: catalytic activity; INVOLVED IN: metabolic process; LOCATED IN: cellular_component unknown; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Glyoxalase/bleomycin resistance protein/dioxygenase (InterPro:IPR004360); BEST Arabidopsis thaliana protein match is: lactoylglutathione lyase family protein / glyoxalase I family protein (TAIR:AT5G41650.1); Has 43 Blast hits to 43 proteins in 18 species: Archae - 0; Bacteria - 12; Metazoa - 0; Fungi - 0; Plants - 24; Viruses - 0; Other Eukaryotes - 7 (source: NCBI BLink). |
| GmaAffx.49521.1.S1_ | 0.08446 | -2.3353 | Glyma11g04870.1 | AT3G49680.1 | Encodes a chloroplast branched-chain amino acid aminotransferase. Complements the yeast leu/iso-leu/val auxotrophy mutant. |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.32380.1.S1_ | 0.09106 | -2.3001 | Glyma11g03390.1 | AT1G62600.1 | flavin-containing monooxygenase family protein / FMO family protein; FUNCTIONS IN: NADP or NADPH binding, monooxygenase activity, FAD binding, flavin-containing monooxygenase activity; INVOLVED IN: biological_process unknown; LOCATED IN: vacuole; CONTAINS InterPro DOMAIN/s: Flavin-containing monooxygenase FMO (InterPro:IPR000960); BEST Arabidopsis thaliana protein match is: flavin-containing monooxygenase family protein / FMO family protein (TAIR:AT1G62620.1); Has 8846 Blast hits to 8478 proteins in 928 species: Archae - 35; Bacteria - 3635; Metazoa - 1030; Fungi - 886; Plants - 455; Viruses - 0; Other Eukaryotes - 2805 (source: NCBI BLink). |
| Gma.6062.2.S1_a_at | 0.03261 | -2.2655 | Glyma19g34890.2 | AT4G14360.2 | dehydration-responsive protein-related; LOCATED IN: Golgi apparatus; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF248, methyltransferase putative (InterPro:IPR004159); BEST Arabidopsis thaliana protein match is: dehydration-responsive protein-related (TAIR:AT3G23300.1); Has 524 Blast hits to 508 proteins in 37 species: Archae - 0; Bacteria - 29; Metazoa - 0; Fungi - 0; Plants - 480; Viruses - 0; Other Eukaryotes - 15 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.5590.1.S1_at | 0.09403 | -2.2327 | Glyma14g35370.1 | AT1G22410.1 | 2-dehydro-3-deoxyphosphoheptonate aldolase, putative / 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase, putative / DAHP synthetase, putative; FUNCTIONS IN: 3-deoxy-7-phosphoheptulonate synthase activity; INVOLVED IN: aromatic amino acid family biosynthetic process; LOCATED IN: membrane; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: DAHP synthetase, class II (InterPro:IPR002480); BEST Arabidopsis thaliana protein match is: DHS1 (3-DEOXY-D-ARABINO-HEPTULOSONATE 7-PHOSPHATE SYNTHASE 1); 3-deoxy-7-phosphoheptulonate synthase (TAIR:AT4G39980.1); Has 3140 Blast hits to 3125 proteins in 395 species: Archae - 0; Bacteria - 662; Metazoa - 0; Fungi - 68; Plants - 129; Viruses - 0; Other Eukaryotes - 2281 (source: NCBI BLink). |
| Gma.6442.2.S1_at | 0.08729 | -2.204 | Glyma13g00450.1 | AT1G52760.1 | Encodes a lysophospholipase 2 (LysoPL2). Involved in tolerance to cadmium-induced oxidative stress. Binds Acyl-CoA-binding protein 2 (ACBP2). |
| Gma.2590.4.S1_at | 0.03377 | -2.1979 | Glyma14g01290.1 | AT5G48300.1 | Encodes the small subunit of ADP-glucose pyrophosphorylase. The small subunit is the catalytic isoform responsible for ADP-glucose pyrophosphorylase activity. The presence of the small subunit is required for large subunit stability. Two isoforms of the small subunit (ApS1 and ApS2) have been described. ApS1 is the major small subunit isoform present in all plant tissues tested. |

| | | | | | |
|---|---|---|---|---|---|
| Gma.4325.1.S1_s_at | 0.08909 | -2.1801 | Glyma08g42070.2 | AT1G06550.1 | enoyl-CoA hydratase/isomerase family protein; FUNCTIONS IN: 3-hydroxyisobutyryl-CoA hydrolase activity, catalytic activity; INVOLVED IN: fatty acid beta-oxidation, metabolic process; LOCATED IN: cellular_component unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Crotonase, core (InterPro:IPR001753); BEST Arabidopsis thaliana protein match is: CHY1 (BETA-HYDROXYISOBUTYRYL-COA HYDROLASE 1); 3-hydroxyisobutyryl-CoA hydrolase (TAIR:AT5G65940.1); Has 18301 Blast hits to 18297 proteins in 1228 species: Archae - 169; Bacteria - 10699; Metazoa - 998; Fungi - 428; Plants - 255; Viruses - 0; Other Eukaryotes - 5752 (source: NCBI BLink). |
| Gma.9592.1.A1_at | 0.04664 | -2.1777 | Glyma20g31250.1 | AT4G09670.1 | oxidoreductase family protein; FUNCTIONS IN: electron carrier activity, oxidoreductase activity, binding, catalytic activity; INVOLVED IN: metabolic process; LOCATED IN: cellular_component unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Oxidoreductase, N-terminal (InterPro:IPR000683), Oxidoreductase, C-terminal (InterPro:IPR004104), NAD(P)-binding (InterPro:IPR016040); BEST Arabidopsis thaliana protein match is: oxidoreductase family protein (TAIR:AT1G34200.1); Has 10490 Blast hits to 10489 proteins in 1179 species: Archae - 137; Bacteria - 6095; Metazoa - 233; Fungi - 386; Plants - 50; Viruses - 0; Other Eukaryotes - 3589 (source: NCBI BLink). |
| Gma.5578.3.S1_at | 0.07148 | -2.1579 | Glyma03g27740.2 | AT2G40890.1 | encodes coumarate 3-hydroxylase (C3H), a P450-dependent monooxygenase. Involved in lignin biosynthesis and flavonoid biosynthesis. Also affects the biosynthesis of coumarins such as scopoletin and scopolin as a branching-out-pathway from the phenylpropanoid acid level. |

| | | | | | |
|---|---|---|---|---|---|
| Gma.5578.3.S1_x_at | 0.03377 | -2.1565 | Glyma03g27740.2 | AT2G40890.1 | encodes coumarate 3-hydroxylase (C3H), a P450-dependent monooxygenase. Involved in lignin biosynthesis and flavonoid biosynthesis. Also affects the biosynthesis of coumarins such as scopoletin and scopolin as a branching-out-pathway from the phenylpropanoid acid level. |
| Gma.5590.1.S1_a_at | 0.04563 | -2.1443 | Glyma14g35370.1 | AT1G22410.1 | 2-dehydro-3-deoxyphosphoheptonate aldolase, putative / 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase, putative / DAHP synthetase, putative; FUNCTIONS IN: 3-deoxy-7-phosphoheptulonate synthase activity; INVOLVED IN: aromatic amino acid family biosynthetic process; LOCATED IN: membrane; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: DAHP synthetase, class II (InterPro:IPR002480); BEST Arabidopsis thaliana protein match is: DHS1 (3-DEOXY-D-ARABINO-HEPTULOSONATE 7-PHOSPHATE SYNTHASE 1); 3-deoxy-7-phosphoheptulonate synthase (TAIR:AT4G39980.1); Has 3140 Blast hits to 3125 proteins in 395 species: Archae - 0; Bacteria - 662; Metazoa - 0; Fungi - 68; Plants - 129; Viruses - 0; Other Eukaryotes - 2281 (source: NCBI BLink). |
| GmaAffx.92716.1.S1_ | 0.05004 | -2.1375 | Glyma09g04280.3 | AT3G06700.3 | 60S ribosomal protein L29 (RPL29A); FUNCTIONS IN: structural constituent of ribosome; INVOLVED IN: translation, ribosome biogenesis; LOCATED IN: cytosolic large ribosomal subunit, ribosome, nucleolus; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Ribosomal protein L29e (InterPro:IPR002673); BEST Arabidopsis thaliana protein match is: 60S ribosomal protein L29 (RPL29B) (TAIR:AT3G06680.2); Has 567 Blast hits to 567 proteins in 179 species: Archae - 0; Bacteria - 0; Metazoa - 345; Fungi - 69; Plants - 75; Viruses - 0; Other Eukaryotes - 78 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.15323.1.S1_at | 0.05763 | -2.1365 | Glyma20g02400.1 | AT2G35120.1 | glycine cleavage system H protein, mitochondrial, putative; FUNCTIONS IN: glycine dehydrogenase (decarboxylating) activity, ATP binding; INVOLVED IN: glycine catabolic process; LOCATED IN: mitochondrion, glycine cleavage complex; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: 2-oxo acid dehydrogenase, lipoyl-binding site (InterPro:IPR003016), Single hybrid motif (InterPro:IPR011053), Glycine cleavage H-protein (InterPro:IPR002930), Glycine cleavage H-protein, subgroup (InterPro:IPR017453); BEST Arabidopsis thaliana protein match is: glycine cleavage system H protein, mitochondrial, putative (TAIR:AT1G32470.1); Has 4762 Blast hits to 4762 proteins in 1230 species: Archae - 82; Bacteria - 2410; Metazoa - 143; Fungi - 86; Plants - 158; Viruses - 0; Other Eukaryotes - 1883 (source: NCBI BLink). |
| GmaAffx.74194.1.S1_ | 0.05226 | -2.1267 | Glyma05g10390.1 | AT3G50560.1 | short-chain dehydrogenase/reductase (SDR) family protein; FUNCTIONS IN: oxidoreductase activity, binding, catalytic activity; INVOLVED IN: metabolic process; LOCATED IN: cellular_component unknown; EXPRESSED IN: 19 plant structures; EXPRESSED DURING: 11 growth stages; CONTAINS InterPro DOMAIN/s: NAD(P)-binding (InterPro:IPR016040), Glucose/ribitol dehydrogenase (InterPro:IPR002347), Short-chain dehydrogenase/reductase SDR (InterPro:IPR002198); BEST Arabidopsis thaliana protein match is: short-chain dehydrogenase/reductase (SDR) family protein (TAIR:AT5G10050.1); Has 18677 Blast hits to 18675 proteins in 1420 species: Archae - 161; Bacteria - 12385; Metazoa - 1272; Fungi - 507; Plants - 217; Viruses - 0; Other Eukaryotes - 4135 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.15480.1.S1_at | 0.09647 | -2.1166 | Glyma13g20010.3 | AT5G02960.1 | 40S ribosomal protein S23 (RPS23B); FUNCTIONS IN: structural constituent of ribosome; INVOLVED IN: translation; LOCATED IN: cytosolic small ribosomal subunit, cytosolic ribosome, ribosome; EXPRESSED IN: guard cell, leaf; CONTAINS InterPro DOMAIN/s: Ribosomal protein S23, eukaryotic/archaeal (InterPro:IPR005680), Ribosomal protein S12/S23 (InterPro:IPR006032), Nucleic acid-binding, OB-fold-like (InterPro:IPR016027), Nucleic acid-binding, OB-fold (InterPro:IPR012340); BEST Arabidopsis thaliana protein match is: 40S ribosomal protein S23 (RPS23A) (TAIR:AT3G09680.1); Has 6124 Blast hits to 6121 proteins in 1817 species: Archae - 180; Bacteria - 2807; Metazoa - 329; Fungi - 198; Plants - 755; Viruses - 0; Other Eukaryotes - 1855 (source: NCBI BLink). |
| GmaAffx.84746.1.A1 | 0.01639 | -2.0904 | Glyma06g05710.1 | AT4G33110.1 | coclaurine N-methyltransferase, putative; FUNCTIONS IN: cyclopropane-fatty-acyl-phospholipid synthase activity, (S)-coclaurine-N-methyltransferase activity; INVOLVED IN: lipid biosynthetic process; LOCATED IN: plasma membrane; EXPRESSED IN: cultured cell; CONTAINS InterPro DOMAIN/s: Cyclopropane-fatty-acyl-phospholipid synthase (InterPro:IPR003333); BEST Arabidopsis thaliana protein match is: coclaurine N-methyltransferase, putative (TAIR:AT4G33120.1); Has 6032 Blast hits to 6030 proteins in 921 species: Archae - 32; Bacteria - 2504; Metazoa - 56; Fungi - 238; Plants - 239; Viruses - 0; Other Eukaryotes - 2963 (source: NCBI BLink). |
| Gma.1314.1.S1_at | 0.06018 | -2.0829 | Glyma13g34290.1 | AT1G78660.2 | The Arabidopsis protein AtGGH1 is a gamma-glutamyl hydrolase cleaving pentaglutamates to yield di- and triglutamates. The enzyme is involved in the tetrahydrofolate metabolism and located to the vacuole. |

| | | | | | |
|---|---|---|---|---|---|
| Gma.12340.1.S1_at | 0.06291 | -2.0758 | Glyma20g29640.1 | AT5G52190.1 | sugar isomerase (SIS) domain-containing protein; FUNCTIONS IN: sugar binding; INVOLVED IN: carbohydrate metabolic process; EXPRESSED IN: 17 plant structures; EXPRESSED DURING: 10 growth stages; CONTAINS InterPro DOMAIN/s: Sugar isomerase (SIS) (InterPro:IPR001347); Has 410 Blast hits to 410 proteins in 146 species: Archae - 96; Bacteria - 262; Metazoa - 0; Fungi - 0; Plants - 15; Viruses - 0; Other Eukaryotes - 37 (source: NCBI BLink). |
| Gma.17484.1.A1_at | 0.05809 | -2.0362 | No Soybean Match Identified, E<10E-30, Perecent ID>95% | AT4G25300.1 | oxidoreductase, 2OG-Fe(II) oxygenase family protein; FUNCTIONS IN: oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen, 2-oxoglutarate as one donor, and incorporation of one atom each of oxygen into both donors, oxidoreductase activity; INVOLVED IN: flavonoid biosynthetic process; LOCATED IN: cellular_component unknown; CONTAINS InterPro DOMAIN/s: 2OG-Fe(II) oxygenase (InterPro:IPR005123); BEST Arabidopsis thaliana protein match is: oxidoreductase, 2OG-Fe(II) oxygenase family protein (TAIR:AT4G25310.1); Has 5835 Blast hits to 5806 proteins in 680 species: Archae - 0; Bacteria - 700; Metazoa - 111; Fungi - 623; Plants - 3093; Viruses - 0; Other Eukaryotes - 1308 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.4529.1.S1_at | 0.05603 | -2.0227 | Glyma09g01050.1 | AT1G12000.1 | pyrophosphate--fructose-6-phosphate 1-phosphotransferase beta subunit, putative / pyrophosphate-dependent 6-phosphofructose-1-kinase, putative; FUNCTIONS IN: diphosphate-fructose-6-phosphate 1-phosphotransferase activity; INVOLVED IN: response to cadmium ion, glycolysis; LOCATED IN: pyrophosphate-dependent phosphofructokinase complex, beta-subunit complex, cell wall, membrane; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Pyrophosphate-dependent phosphofructokinase PfpB (InterPro:IPR011183), Phosphofructokinase (InterPro:IPR000023); BEST Arabidopsis thaliana protein match is: MEE51 (maternal effect embryo arrest 51); diphosphate-fructose-6-phosphate 1-phosphotransferase (TAIR:AT4G04040.1); Has 3585 Blast hits to 3515 proteins in 1006 species: Archae - 20; Bacteria - 2321; Metazoa - 52; Fungi - 90; Plants - 237; Viruses - 2; Other Eukaryotes - 863 (source: NCBI BLink). |
| Gma.7596.1.S1_at | 0.06613 | -2.0214 | Glyma18g15260.1 | AT2G46210.1 | delta-8 sphingolipid desaturase, putative; FUNCTIONS IN: oxidoreductase activity, oxidoreductase activity, acting on paired donors, with oxidation of a pair of donors resulting in the reduction of molecular oxygen to two molecules of water, iron ion binding, heme binding; INVOLVED IN: fatty acid biosynthetic process, lipid metabolic process; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Fatty acid desaturase, type 1 (InterPro:IPR005804), Cytochrome b5 (InterPro:IPR001199), Fatty acid/sphingolipid desaturase (InterPro:IPR012171); BEST Arabidopsis thaliana protein match is: delta-8 sphingolipid desaturase (SLD1) (TAIR:AT3G61580.1); Has 4212 Blast hits to 4149 proteins in 626 species: Archae - 1; Bacteria - 637; Metazoa - 966; Fungi - 1090; Plants - 641; Viruses - 2; Other Eukaryotes - 875 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.17178.1.S1_s_at | 0.06366 | -2.0168 | Glyma01g37900.1 | AT5G66530.1 | aldose 1-epimerase family protein; FUNCTIONS IN: carbohydrate binding, isomerase activity, aldose 1-epimerase activity, catalytic activity; INVOLVED IN: galactose metabolic process, carbohydrate metabolic process; LOCATED IN: apoplast, chloroplast, chloroplast stroma; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Glycoside hydrolase-type carbohydrate-binding (InterPro:IPR011013), Aldose 1-epimerase (InterPro:IPR008183), Glycoside hydrolase-type carbohydrate-binding, subgroup (InterPro:IPR014718); BEST Arabidopsis thaliana protein match is: aldose 1-epimerase family protein (TAIR:AT5G57330.1); Has 1296 Blast hits to 1295 proteins in 534 species: Archae - 0; Bacteria - 866; Metazoa - 40; Fungi - 84; Plants - 134; Viruses - 0; Other Eukaryotes - 172 (source: NCBI BLink). |
| GmaAffx.93559.1.S1_ | 0.0958 | -2.0009 | Glyma04g11550.1 | AT1G36160.1 | Encodes acetyl-CoA carboxylase. Mutant displays uncoordinated cell divisions which are enhanced by cytokinins. Mutant also has aberrant organization of the apical region in the embryo and abnormal root and shoot development. Essential for very long chain fatty acid elongation. |
| GmaAffx.16986.1.A1 | 0.09942 | -1.9982 | No Soybean Match Identified, E<10E-30, Perecent ID>95% | AT5G54160.1 | A caffeic acid/5-hydroxyferulic acid O-methyltransferase. Interacts with 14-4-3 proteins in yeast 2 hybrid assay. AtOMT1 (At5g54160) encodes a flavonol 3'-O-methyltransferase that is highly active towards quercetin and myricetin. The substrate specificity identifies the enzyme as flavonol 3'-methyltransferase which replaces the former annotation of the gene to encode a caffeic acid/5-hydroxyferulic acid O-methyltransferase |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.36917.1.S1_ | 0.07344 | -1.96 | Glyma03g32510.2 | AT2G20340.1 | tyrosine decarboxylase, putative; FUNCTIONS IN: pyridoxal phosphate binding, carboxy-lyase activity, catalytic activity, tyrosine decarboxylase activity; INVOLVED IN: amino acid metabolic process, response to wounding; LOCATED IN: cellular_component unknown; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Aromatic-L-amino-acid decarboxylase (InterPro:IPR010977), Pyridoxal phosphate-dependent transferase, major region (InterPro:IPR015424), Pyridoxal phosphate-dependent decarboxylase (InterPro:IPR002129), Pyridoxal phosphate-dependent transferase, major region, subdomain 1 (InterPro:IPR015421), Pyridoxal phosphate-dependent transferase, major region, subdomain 2 (InterPro:IPR015422); BEST Arabidopsis thaliana protein match is: tyrosine decarboxylase, putative (TAIR:AT4G28680.1); Has 3991 Blast hits to 3974 proteins in 1223 species: Archae - 53; Bacteria - 1092; Metazoa - 1910; Fungi - 168; Plants - 143; Viruses - 5; Other Eukaryotes - 620 (source: NCBI BLink). |
| Gma.11134.1.A1_at | 0.07119 | -1.9423 | Glyma08g41630.1 | AT1G06690.1 | aldo/keto reductase family protein; FUNCTIONS IN: oxidoreductase activity, aldo-keto reductase activity, ATPase activity, ATP binding; INVOLVED IN: oxidation reduction; LOCATED IN: chloroplast thylakoid membrane, chloroplast, plastoglobule, chloroplast envelope; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Aldo/keto reductase (InterPro:IPR001395), Aldo/keto reductase, conserved site (InterPro:IPR018170), ABC transporter, conserved site (InterPro:IPR017871); BEST Arabidopsis thaliana protein match is: aldo/keto reductase family protein (TAIR:AT5G53580.1); Has 18029 Blast hits to 18016 proteins in 1454 species: Archae - 258; Bacteria - 9842; Metazoa - 1647; Fungi - 1390; Plants - 743; Viruses - 0; Other Eukaryotes - 4149 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.6059.1.S1_at | 0.0585 | -1.933 | Glyma09g36680.3 | AT4G25280.1 | adenylate kinase family protein; FUNCTIONS IN: nucleobase, nucleoside, nucleotide kinase activity, nucleotide kinase activity, phosphotransferase activity, phosphate group as acceptor, ATP binding; INVOLVED IN: nucleobase, nucleoside, nucleotide and nucleic acid metabolic process; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: UMP-CMP kinase (InterPro:IPR006266), Adenylate kinase (InterPro:IPR000850); BEST Arabidopsis thaliana protein match is: PYR6; cytidylate kinase/ uridylate kinase (TAIR:AT5G26667.3); Has 8572 Blast hits to 8424 proteins in 1844 species: Archae - 61; Bacteria - 4360; Metazoa - 1053; Fungi - 294; Plants - 243; Viruses - 0; Other Eukaryotes - 2561 (source: NCBI BLink). |
| GmaAffx.33989.1.A1 | 0.0577 | -1.9295 | Glyma10g03930.1 | AT4G19540.1 | IND1(IRON-SULFUR PROTEIN REQUIRED FOR NADH DEHYDROGENASE)-LIKE (INDL); LOCATED IN: mitochondrion; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 13 growth stages; BEST Arabidopsis thaliana protein match is: HCF101 (HIGH-CHLOROPHYLL-FLUORESCENCE 101); ATP binding (TAIR:AT3G24430.1); Has 9217 Blast hits to 9199 proteins in 1435 species: Archae - 352; Bacteria - 4208; Metazoa - 319; Fungi - 279; Plants - 117; Viruses - 0; Other Eukaryotes - 3942 (source: NCBI BLink). |
| Gma.890.1.S1_at | 0.05544 | -1.9252 | Glyma15g03120.1 | AT4G29840.1 | threonine synthase |
| Gma.2222.1.S1_at | 0.04522 | -1.9059 | Glyma04g01170.3 | AT4G39120.1 | Encodes a myo-inositol monophosphatase IMPL2 (myo-Inositol monophosphatase like 2). |
| GmaAffx.38611.1.A1 | 0.05004 | -1.8825 | Glyma10g03410.2 | AT2G32010.1 | Encodes an inositol polyphosphate 5'-phosphatase (5PTase). Mediating phosphoinositide signaling. Involved in establishment of foliar vein patterns. |
| GmaAffx.89562.1.A1 | 0.09586 | -1.879 | Glyma08g11480.1 | AT4G13940.1 | Encodes a S-adenosyl-L-homocysteine hydrolase required for DNA methylation-dependent gene silencing. |

| | | | | | |
|---|---|---|---|---|---|
| Gma.17941.1.S1_at | 0.07789 | -1.8645 | Glyma03g40110.1 | AT3G52580.1 | 40S ribosomal protein S14 (RPS14C); FUNCTIONS IN: structural constituent of ribosome; INVOLVED IN: translation; LOCATED IN: cytosolic small ribosomal subunit; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Ribosomal protein S11 (InterPro:IPR001971), Ribosomal S11, conserved site (InterPro:IPR018102); BEST Arabidopsis thaliana protein match is: 40S ribosomal protein S14 (RPS14B) (TAIR:AT3G11510.1); Has 6258 Blast hits to 6258 proteins in 1752 species: Archae - 169; Bacteria - 2837; Metazoa - 489; Fungi - 109; Plants - 512; Viruses - 0; Other Eukaryotes - 2142 (source: NCBI BLink). |
| Gma.16508.1.S1_at | 0.02727 | -1.8594 | No Soybean Match Identified, E<10E-30, Perecent ID>95% | AT4G28706.3 | pfkB-type carbohydrate kinase family protein; FUNCTIONS IN: kinase activity, ribokinase activity; INVOLVED IN: D-ribose metabolic process; LOCATED IN: chloroplast; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Carbohydrate/purine kinase (InterPro:IPR011611), Ribokinase (InterPro:IPR002139), Carbohydrate/puine kinase, PfkB, conserved site (InterPro:IPR002173); BEST Arabidopsis thaliana protein match is: pfkB-type carbohydrate kinase family protein (TAIR:AT5G43910.2); Has 4369 Blast hits to 4279 proteins in 956 species: Archae - 120; Bacteria - 2818; Metazoa - 278; Fungi - 82; Plants - 158; Viruses - 0; Other Eukaryotes - 913 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.15478.1.S1_x_at | 0.05394 | -1.8496 | Glyma1337s00200.1 | AT4G01850.2 | S-ADENOSYLMETHIONINE SYNTHETASE 2 (SAM-2); FUNCTIONS IN: methionine adenosyltransferase activity; INVOLVED IN: one-carbon compound metabolic process, S-adenosylmethionine biosynthetic process; EXPRESSED IN: 27 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: S-adenosylmethionine synthetase (InterPro:IPR002133); BEST Arabidopsis thaliana protein match is: SAM1 (S-ADENOSYLMETHIONINE SYNTHETASE 1); methionine adenosyltransferase (TAIR:AT1G02500.2); Has 8027 Blast hits to 8026 proteins in 1640 species: Archae - 6; Bacteria - 2999; Metazoa - 309; Fungi - 111; Plants - 502; Viruses - 0; Other Eukaryotes - 4100 (source: NCBI BLink). |
| Gma.6843.1.S1_at | 0.07755 | -1.8342 | Glyma15g04500.2 | AT2G28760.3 | UDP-XYL SYNTHASE 6 (UXS6); FUNCTIONS IN: coenzyme binding, binding, catalytic activity; INVOLVED IN: cellular metabolic process, nucleotide-sugar metabolic process, metabolic process; LOCATED IN: plasma membrane; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: NAD-dependent epimerase/dehydratase (InterPro:IPR001509), NAD(P)-binding (InterPro:IPR016040); BEST Arabidopsis thaliana protein match is: UXS5; UDP-glucuronate decarboxylase/ catalytic (TAIR:AT3G46440.2); Has 32193 Blast hits to 32159 proteins in 1749 species: Archae - 518; Bacteria - 14743; Metazoa - 694; Fungi - 281; Plants - 822; Viruses - 72; Other Eukaryotes - 15063 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.4904.1.S1_at | 0.0687 | -1.8164 | Glyma04g10100.4 | AT5G10920.1 | argininosuccinate lyase, putative / arginosuccinase, putative; FUNCTIONS IN: argininosuccinate lyase activity, catalytic activity; INVOLVED IN: arginine biosynthetic process via ornithine, arginine biosynthetic process; LOCATED IN: chloroplast, chloroplast stroma; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Argininosuccinate lyase (InterPro:IPR009049), L-Aspartase-like (InterPro:IPR008948), Delta crystallin (InterPro:IPR003031), Fumarate lyase (InterPro:IPR000362); Has 10040 Blast hits to 10035 proteins in 1495 species: Archae - 255; Bacteria - 5320; Metazoa - 254; Fungi - 173; Plants - 41; Viruses - 0; Other Eukaryotes - 3997 (source: NCBI BLink). |
| Gma.1410.1.S1_at | 0.06659 | -1.8158 | Glyma08g23860.1 | AT1G17745.1 | encodes a 3-Phosphoglycerate dehydrogenase |
| Gma.3220.1.A1_at | 0.0577 | -1.8035 | Glyma14g00340.1 | AT2G17265.1 | Encodes a homoserine kinase (HSK) which produces O-phospho-L-homoserine (HserP), a compound at the branching point of methionine and threonine biosynthesis. HSK is found in the stromal fraction of chloroplasts. Mutants are susceptible to downy mildew fungus Hyaloperonospora parasitica. |
| GmaAffx.1681.1.A1_ | 0.06238 | -1.8006 | Glyma07g40340.4 | AT1G57770.1 | amine oxidase family; FUNCTIONS IN: oxidoreductase activity; LOCATED IN: chloroplast; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: FAD dependent oxidoreductase (InterPro:IPR006076); BEST Arabidopsis thaliana protein match is: CRTISO (CAROTENOID ISOMERASE); carotenoid isomerase (TAIR:AT1G06820.1); Has 4784 Blast hits to 4705 proteins in 573 species: Archae - 90; Bacteria - 1781; Metazoa - 346; Fungi - 53; Plants - 173; Viruses - 0; Other Eukaryotes - 2341 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.15027.1.S1_at | 0.09251 | -1.7976 | Glyma04g05420.2 | AT4G33010.1 | Arabidopsis thaliana glycine decarboxylase P-protein 1 (AtGLDP1); FUNCTIONS IN: glycine dehydrogenase (decarboxylating) activity, pyridoxal phosphate binding, catalytic activity; INVOLVED IN: glycine catabolic process, glycine decarboxylation via glycine cleavage system; LOCATED IN: mitochondrion, apoplast, glycine cleavage complex, chloroplast, chloroplast envelope; EXPRESSED IN: 31 plant structures; EXPRESSED DURING: 16 growth stages; CONTAINS InterPro DOMAIN/s: Pyridoxal phosphate-dependent transferase, major region (InterPro:IPR015424), Glycine cleavage system P-protein (InterPro:IPR003437), Pyridoxal phosphate-dependent transferase, major region, subdomain 1 (InterPro:IPR015421); BEST Arabidopsis thaliana protein match is: AtGLDP2 (Arabidopsis thaliana glycine decarboxylase P-protein 2); ATP binding / glycine dehydrogenase (decarboxylating) (TAIR:AT2G26080.1); Has 9787 Blast hits to 8954 proteins in 1113 species: Archae - 131; Bacteria - 2880; Metazoa - 122; Fungi - 158; Plants - 73; Viruses - 0; Other Eukaryotes - 6423 (source: NCBI BLink). |
| Gma.11705.1.A1_at | 0.07428 | -1.795 | No Soybean Match Identified, E<10E-30, Perecent ID>95% | AT1G08630.4 | Encodes a threonine aldolase, involved in threonine degradation to glycine. Primarily expressed in seeds and seedlings. |
| GmaAffx.83218.1.S1_ | 0.06659 | -1.7844 | Glyma09g05600.4 | AT4G20070.1 | The gene encoding Arabidopsis thaliana Allantoate Amidohydrolase (AtAAH)which catalyzes the allantoate deiminase reaction (EC 3.5.3.9)is expressed in all parts of the plant being consistent with a function in purine turnover in Arabidopsis. |
| Gma.8431.1.S1_at | 0.03908 | -1.7844 | Glyma10g29780.1 | AT1G09830.1 | glycinamide ribonucleotide synthetase (GAR synthetase) that catalyzes the conversion of phosphoribosyl amine to phosphoribosyl glycineamide |
| GmaAffx.93605.1.S1_ | 0.09725 | -1.7765 | No Soybean Match Identified, E<10E-30, Perecent ID>95% | AT1G23800.1 | Encodes a mitochondrial aldehyde dehydrogenase; nuclear gene for mitochondrial product. |

| | | | | | |
|---|---|---|---|---|---|
| Gma.1355.2.S1_s_at | 0.06238 | -1.7636 | Glyma18g05600.1 | AT3G54210.1 | ribosomal protein L17 family protein; FUNCTIONS IN: structural constituent of ribosome; INVOLVED IN: translation; LOCATED IN: ribosome, chloroplast; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Ribosomal protein L17 (InterPro:IPR000456); BEST Arabidopsis thaliana protein match is: ribosomal protein L17 family protein (TAIR:AT5G09770.1); Has 5385 Blast hits to 5385 proteins in 1530 species: Archae - 0; Bacteria - 3017; Metazoa - 102; Fungi - 83; Plants - 73; Viruses - 0; Other Eukaryotes - 2110 (source: NCBI BLink). |
| Gma.10630.1.S1_at | 0.07586 | -1.7547 | No Soybean Match Identified, E<10E-30, Perecent ID>95% | AT2G39290.1 | Encodes a phosphatidylglycerolphosphate synthase 2C which is dual-targeted into chloroplasts and mitochondria. Mutant plants have mutant chloroplasts but normal mitochondria. |
| Gma.8251.1.S1_a_at | 0.04002 | -1.754 | Glyma10g40870.3 | AT3G19450.1 | Encodes a catalytically active cinnamyl alcohol dehydrogenase which uses p-coumaryl aldehyde as a preferred substrate. It can also use caffeyl, coniferyl and d-hydroxyconiferyl aldehydes as substrates. |
| Gma.2041.1.S1_at | 0.07427 | -1.7503 | Glyma11g33700.1 | AT5G16710.1 | The protein undergoes thiolation following treatment with the oxidant tert-butylhydroperoxide. |
| Gma.881.1.S1_at | 0.03697 | -1.7324 | Glyma12g30660.1 | AT2G27820.1 | Encodes a plastid-localized arogenate dehydratase involved in phenylalanine biosynthesis. Not less than six genes encoding ADT were identiﬁed in the Arabidopsis genome: ADT1 [At1g11790]; ADT2 [At3g07630]; ADT3 [At2g27820]; ADT4 [At3g44720]; ADT5 [At5g22630]; and ADT6 [At1g08250]. |

| | | | | | |
|---|---|---|---|---|---|
| Gma.3087.1.S1_at | 0.06291 | -1.727 | Glyma06g06480.1 | AT3G23780.1 | This gene encodes the second largest, catalytic subunit of the nuclear DNA-dependent RNA polymerase IV (aka RNA polymerase D). The NRPD2 protein is found at nuclear foci that overlap or are adjacent to chromocentromeres but are not fully coincident with chromocentromeres. The loss of NRPD2 leads to the loss of cytosine methylation at pericentromeric 5S genes and AtSN1 retroelements but has no discernible effect on centromere repeat methylation. This suggests that Pol IV primarily affects facultative heterochromatin rather than constitutive heterochromatin. |
| Gma.15478.2.S1_at | 0.06925 | -1.7244 | Glyma17g04340.1 | AT4G01850.2 | S-ADENOSYLMETHIONINE SYNTHETASE 2 (SAM-2); FUNCTIONS IN: methionine adenosyltransferase activity; INVOLVED IN: one-carbon compound metabolic process, S-adenosylmethionine biosynthetic process; EXPRESSED IN: 27 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: S-adenosylmethionine synthetase (InterPro:IPR002133); BEST Arabidopsis thaliana protein match is: SAM1 (S-ADENOSYLMETHIONINE SYNTHETASE 1); methionine adenosyltransferase (TAIR:AT1G02500.2); Has 8027 Blast hits to 8026 proteins in 1640 species: Archae - 6; Bacteria - 2999; Metazoa - 309; Fungi - 111; Plants - 502; Viruses - 0; Other Eukaryotes - 4100 (source: NCBI BLink). |
| Gma.2105.1.S1_at | 0.06659 | -1.7192 | Glyma09g30370.3 | AT5G37600.1 | encodes a cytosolic glutamine synthetase, the enzyme has high affinity with substrate ammonium |

| | | | | | |
|---|---|---|---|---|---|
| Gma.3806.1.S1_at | 0.08166 | -1.7072 | Glyma19g45120.2 | AT2G46890.1 | oxidoreductase, acting on the CH-CH group of donors; FUNCTIONS IN: oxidoreductase activity, acting on the CH-CH group of donors; INVOLVED IN: lipid metabolic process; LOCATED IN: endomembrane system, integral to membrane, cytoplasm; CONTAINS InterPro DOMAIN/s: 3-oxo-5-alpha-steroid 4-dehydrogenase, C-terminal (InterPro:IPR001104), Protein of unknown function DUF1295 (InterPro:IPR010721); BEST Arabidopsis thaliana protein match is: oxidoreductase, acting on the CH-CH group of donors (TAIR:AT1G18180.1); Has 1542 Blast hits to 1542 proteins in 207 species: Archae - 2; Bacteria - 258; Metazoa - 69; Fungi - 95; Plants - 57; Viruses - 0; Other Eukaryotes - 1061 (source: NCBI BLink). |
| Gma.16076.1.A1_at | 0.06637 | -1.6976 | Glyma08g16970.1 | AT1G05590.1 | Encodes a protein with β-hexosaminidase activity (the enzyme is active with p-nitrophenyl-β-N-acetylglucosaminide as substrate but displayed only a minor activity toward p-nitrophenyl-β-N-acetylgalactosaminide). Chitotriose-PA was digested almost completely overnight by a 50% ammonium sulfate fraction of a supernatant yeast expressing AtHEX3. |

| | | | | | |
|---|---|---|---|---|---|
| Gma.17648.1.S1_at | 0.06847 | -1.689 | Glyma18g20800.1 | AT2G42490.1 | copper amine oxidase, putative; FUNCTIONS IN: quinone binding, amine oxidase activity, copper ion binding; INVOLVED IN: cellular amine metabolic process; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Copper amine oxidase, N-terminal region (InterPro:IPR016182), Copper amine oxidase, N2-terminal (InterPro:IPR015800), Copper amine oxidase, N2/N3-terminal (InterPro:IPR015801), Copper amine oxidase, N3-terminal (InterPro:IPR015802), Copper amine oxidase (InterPro:IPR000269), Copper amine oxidase, C-terminal (InterPro:IPR015798); BEST Arabidopsis thaliana protein match is: copper amine oxidase, putative (TAIR:AT3G43670.1); Has 1087 Blast hits to 1084 proteins in 164 species: Archae - 4; Bacteria - 240; Metazoa - 211; Fungi - 291; Plants - 130; Viruses - 0; Other Eukaryotes - 211 (source: NCBI BLink). |
| Gma.6954.2.S1_at | 0.04759 | -1.6885 | No Soybean Match Identified, E<10E-30, Perecent ID>95% | AT5G17710.2 | embryo defective 1241 (EMB1241); FUNCTIONS IN: protein binding, adenyl-nucleotide exchange factor activity, chaperone binding, protein homodimerization activity; INVOLVED IN: embryonic development ending in seed dormancy, protein folding; LOCATED IN: thylakoid, chloroplast; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: GrpE nucleotide exchange factor (InterPro:IPR000740), GrpE nucleotide exchange factor, coiled-coil (InterPro:IPR013805), GrpE nucleotide exchange factor, head (InterPro:IPR009012); BEST Arabidopsis thaliana protein match is: co-chaperone grpE family protein (TAIR:AT1G36390.2); Has 5868 Blast hits to 5837 proteins in 1527 species: Archae - 92; Bacteria - 2831; Metazoa - 204; Fungi - 113; Plants - 94; Viruses - 7; Other Eukaryotes - 2527 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.12507.1.S1_at | 0.05948 | -1.6831 | Glyma09g00980.1 | AT2G20060.1 | ribosomal protein L4 family protein; FUNCTIONS IN: structural constituent of ribosome; INVOLVED IN: translation; LOCATED IN: ribosome, intracellular, chloroplast, large ribosomal subunit; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Ribosomal protein L4 (InterPro:IPR015498), Ribosomal protein L4/L1e, bacterial-type (InterPro:IPR013005), Ribosomal protein L4/L1e (InterPro:IPR002136); BEST Arabidopsis thaliana protein match is: RPL4; poly(U) binding / structural constituent of ribosome (TAIR:AT1G07320.4); Has 5451 Blast hits to 5451 proteins in 1497 species: Archae - 63; Bacteria - 2944; Metazoa - 100; Fungi - 82; Plants - 60; Viruses - 0; Other Eukaryotes - 2202 (source: NCBI BLink). |
| Gma.3953.1.S1_at | 0.07148 | -1.682 | Glyma03g36550.2 | AT3G58560.1 | endonuclease/exonuclease/phosphatase family protein; FUNCTIONS IN: hydrolase activity; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Endonuclease/exonuclease/phosphatase (InterPro:IPR005135); BEST Arabidopsis thaliana protein match is: hydrolase (TAIR:AT3G58580.1); Has 965 Blast hits to 938 proteins in 154 species: Archae - 0; Bacteria - 7; Metazoa - 452; Fungi - 163; Plants - 175; Viruses - 0; Other Eukaryotes - 168 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.91087.1.S1_ | 0.08973 | -1.6812 | Glyma08g02610.1 | AT1G64390.1 | Arabidopsis thaliana glycosyl hydrolase 9C2 (AtGH9C2); FUNCTIONS IN: carbohydrate binding, hydrolase activity, hydrolyzing O-glycosyl compounds, catalytic activity; INVOLVED IN: carbohydrate metabolic process; LOCATED IN: endomembrane system; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Six-hairpin glycosidase (InterPro:IPR012341), Carbohydrate-binding (InterPro:IPR008965), Glycoside hydrolase, family 9, active site (InterPro:IPR018221), Six-hairpin glycosidase-like (InterPro:IPR008928), Glycoside hydrolase, family 9 (InterPro:IPR001701); BEST Arabidopsis thaliana protein match is: AtGH9C3 (Arabidopsis thaliana glycosyl hydrolase 9C3); carbohydrate binding / catalytic/ hydrolase, hydrolyzing O-glycosyl compounds (TAIR:AT4G11050.1); Has 1169 Blast hits to 1156 proteins in 190 species: Archae - 0; Bacteria - 327; Metazoa - 138; Fungi - 14; Plants - 625; Viruses - 0; Other Eukaryotes - 65 (source: NCBI BLink). |
| Gma.2960.1.S1_at | 0.07755 | -1.6807 | Glyma05g04300.1 | AT1G12230.1 | transaldolase, putative; FUNCTIONS IN: transaldolase activity, catalytic activity; INVOLVED IN: glucose catabolic process to lactate and acetate, 5-phosphoribose 1-diphosphate biosynthetic process, carbohydrate metabolic process, metabolic process, pentose-phosphate shunt, non-oxidative branch; LOCATED IN: chloroplast; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Aldolase-type TIM barrel (InterPro:IPR013785), Transaldolase (InterPro:IPR001585); Has 4687 Blast hits to 4687 proteins in 1248 species: Archae - 40; Bacteria - 2981; Metazoa - 141; Fungi - 151; Plants - 60; Viruses - 7; Other Eukaryotes - 1307 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.92501.1.S1_ | 0.06963 | -1.6797 | Glyma10g00920.2 | AT5G43330.1 | malate dehydrogenase, cytosolic, putative; FUNCTIONS IN: in 6 functions; INVOLVED IN: cellular carbohydrate metabolic process, glycolysis, malate metabolic process, carbohydrate metabolic process, metabolic process; LOCATED IN: plasma membrane; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Malate dehydrogenase, NAD-dependent, cytosolic (InterPro:IPR011274), Lactate/malate dehydrogenase (InterPro:IPR001236), Malate dehydrogenase, NAD or NADP (InterPro:IPR010945), NAD(P)-binding (InterPro:IPR016040), L-lactate/malate dehydrogenase (InterPro:IPR001557), Malate dehydrogenase, active site (InterPro:IPR001252), Lactate dehydrogenase/glycoside hydrolase, family 4, C-terminal (InterPro:IPR015955); BEST Arabidopsis thaliana protein match is: malate dehydrogenase, cytosolic, putative (TAIR:AT1G04410.1); Has 8255 Blast hits to 8253 proteins in 1762 species: Archae - 116; Bacteria - 4254; Metazoa - 1034; Fungi - 184; Plants - 466; Viruses - 0; Other Eukaryotes - 2201 (source: NCBI BLink). |
| GmaAffx.58488.1.S1_ | 0.08317 | -1.6792 | Glyma18g47650.1 | AT1G65295.1 | unknown protein; LOCATED IN: endomembrane system; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT5G01015.1); Has 55 Blast hits to 55 proteins in 8 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 55; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| Gma.11154.4.S1_at | 0.06659 | -1.6788 | Glyma12g01690.1 | AT1G64970.1 | gamma-tocopherol methyltransferase (g-TMT) mRNA, nuclear; mutant has Deficient in alpha and beta tocopherol; Accumulates gamma tocopherol in leaves |
| Gma.1138.1.S1_at | 0.1 | -1.6767 | Glyma13g36700.1 | AT4G21470.1 | Bifunctional enzyme that catalyzes hydrolysis of FMN to riboflavin, and phosphorylation of riboflavin to FMN. |
| GmaAffx.36650.1.S1_ | 0.08139 | -1.6751 | Glyma05g29030.1 | AT1G55320.1 | Encodes a protein with similarity to acyl activating enzymes. AAE18 is localized to the peroxisome where it may be involved in metabolism of auxin precursors to active auxins. |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.37677.1.S1_ | 0.09965 | -1.6691 | Glyma01g45020.1 | AT1G47480.1 | hydrolase; FUNCTIONS IN: hydrolase activity; INVOLVED IN: metabolic process; LOCATED IN: cellular_component unknown; EXPRESSED IN: stem, hypocotyl, flower, root; EXPRESSED DURING: petal differentiation and expansion stage; CONTAINS InterPro DOMAIN/s: Lipase, GDXG, active site (InterPro:IPR002168), Alpha/beta hydrolase fold-3 (InterPro:IPR013094); BEST Arabidopsis thaliana protein match is: ATCXE13 (ARABIDOPSIS THALIANA CARBOXYESTERASE 13); hydrolase (TAIR:AT3G48700.1); Has 5187 Blast hits to 5177 proteins in 806 species: Archae - 47; Bacteria - 2656; Metazoa - 445; Fungi - 493; Plants - 650; Viruses - 3; Other Eukaryotes - 893 (source: NCBI BLink). |
| Gma.2354.2.S1_at | 0.09148 | -1.6676 | Glyma20g38950.3 | AT1G74040.1 | Encodes an active Arabidopsis isopropylmalate synthase IPMS2. Involved in leucine biosynthesis. Do not participate in the chain elongation of glucosinolates. Expressed constitutively throughout the plant. Loss of IPMS2 can be compensated by a second isopropylmalate synthase gene IPMS1 (At1g18500). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.91593.1.S1 | 0.09399 | -1.6666 | Glyma13g41960.1 | AT3G59480.1 | pfkB-type carbohydrate kinase family protein; FUNCTIONS IN: kinase activity, ribokinase activity; INVOLVED IN: D-ribose metabolic process, acetate fermentation, sucrose biosynthetic process, sucrose catabolic process, using beta-fructofuranosidase; LOCATED IN: cellular_component unknown; EXPRESSED IN: 9 plant structures; EXPRESSED DURING: 4 anthesis, C globular stage, petal differentiation and expansion stage; CONTAINS InterPro DOMAIN/s: Carbohydrate/purine kinase (InterPro:IPR011611), Ribokinase (InterPro:IPR002139), Carbohydrate/puine kinase, PfkB, conserved site (InterPro:IPR002173); BEST Arabidopsis thaliana protein match is: pfkB-type carbohydrate kinase family protein (TAIR:AT2G31390.1); Has 12101 Blast hits to 12099 proteins in 1317 species: Archae - 211; Bacteria - 7841; Metazoa - 113; Fungi - 101; Plants - 260; Viruses - 0; Other Eukaryotes - 3575 (source: NCBI BLink). |
| Gma.13081.1.S1_at | 0.05811 | -1.6643 | Glyma11g13720.2 | AT2G31440.1 | protein binding; FUNCTIONS IN: protein binding; INVOLVED IN: positive regulation of catalytic activity, protein processing; LOCATED IN: integral to membrane; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Aph-1 (InterPro:IPR009294); Has 215 Blast hits to 211 proteins in 56 species: Archae - 0; Bacteria - 0; Metazoa - 186; Fungi - 0; Plants - 17; Viruses - 0; Other Eukaryotes - 12 (source: NCBI BLink). |
| Gma.4375.1.S1_s_at | 0.05105 | -1.6642 | Glyma18g04660.1 | AT5G37600.1 | encodes a cytosolic glutamine synthetase, the enzyme has high affinity with substrate ammonium |
| Gma.2635.1.A1_s_at | 0.05603 | -1.6585 | Glyma17g05290.1 | AT2G27820.1 | Encodes a plastid-localized arogenate dehydratase involved in phenylalanine biosynthesis. Not less than six genes encoding ADT were identiﬁed in the Arabidopsis genome: ADT1 [At1g11790]; ADT2 [At3g07630]; ADT3 [At2g27820]; ADT4 [At3g44720]; ADT5 [At5g22630]; and ADT6 [At1g08250]. |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.68561.1.S1_ | 0.07385 | -1.6582 | Glyma01g44350.2 | AT5G16800.2 | GCN5-related N-acetyltransferase (GNAT) family protein; FUNCTIONS IN: N-acetyltransferase activity; INVOLVED IN: metabolic process; LOCATED IN: cellular_component unknown; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: GCN5-related N-acetyltransferase (InterPro:IPR000182), Acyl-CoA N-acyltransferase (InterPro:IPR016181); BEST Arabidopsis thaliana protein match is: GCN5-related N-acetyltransferase (GNAT) family protein (TAIR:AT3G02980.1); Has 653 Blast hits to 653 proteins in 208 species: Archae - 49; Bacteria - 129; Metazoa - 223; Fungi - 49; Plants - 61; Viruses - 2; Other Eukaryotes - 140 (source: NCBI BLink). |
| Gma.13247.1.S1_at | 0.09683 | -1.6563 | Glyma05g36500.2 | AT2G07180.1 | protein kinase, putative; FUNCTIONS IN: protein serine/threonine kinase activity, protein kinase activity, kinase activity, ATP binding; INVOLVED IN: protein amino acid phosphorylation, N-terminal protein myristoylation; LOCATED IN: cellular_component unknown; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Protein kinase, ATP binding site (InterPro:IPR017441), Protein kinase, core (InterPro:IPR000719), Serine/threonine protein kinase-related (InterPro:IPR017442), Protein kinase-like (InterPro:IPR011009), Serine/threonine protein kinase, active site (InterPro:IPR008271); BEST Arabidopsis thaliana protein match is: protein kinase family protein (TAIR:AT5G01020.1); Has 83377 Blast hits to 82332 proteins in 3112 species: Archae - 42; Bacteria - 7499; Metazoa - 37069; Fungi - 6338; Plants - 18225; Viruses - 294; Other Eukaryotes - 13910 (source: NCBI BLink). |
| Gma.7812.1.S1_at | 0.02679 | -1.6481 | Glyma09g35670.1 | AT2G16570.1 | Amidophosphoribosyltransferase (ATase: EC 2.4.2.14) is a key enzyme in the pathway of purine nucleotide biosynthesis |
| Gma.1111.1.S1_x_at | 0.06449 | -1.6399 | Glyma18g17260.1 | AT1G59900.1 | encodes the e1 alpha subunit of the pyruvate dehydrogenase complex (PDC) |

| Gma.3419.1.S1_at | 0.0196 | -1.6396 | Glyma20g24990.1 | AT3G25920.1 | encodes a plastid ribosomal protein CL15, a constituent of the large subunit of the ribosomal complex |
|---|---|---|---|---|---|
| GmaAffx.44227.1.S1_ | 0.06304 | -1.6368 | Glyma02g43980.3 | AT1G16740.1 | ribosomal protein L20 family protein; FUNCTIONS IN: structural constituent of ribosome, RNA binding; INVOLVED IN: translation, ribosome biogenesis; LOCATED IN: ribosome, intracellular; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Ribosomal protein L20, bacterial-type (InterPro:IPR005812), Ribosomal protein L20 (InterPro:IPR005813); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:ATCG00660.1); Has 5711 Blast hits to 5711 proteins in 1648 species: Archae - 0; Bacteria - 2947; Metazoa - 105; Fungi - 0; Plants - 471; Viruses - 0; Other Eukaryotes - 2188 (source: NCBI BLink). |
| GmaAffx.92700.1.S1_ | 0.07385 | -1.6305 | Glyma05g37170.1 | AT5G41670.2 | 6-phosphogluconate dehydrogenase family protein; FUNCTIONS IN: in 6 functions; INVOLVED IN: response to fructose stimulus, response to cadmium ion, response to sucrose stimulus, response to glucose stimulus; LOCATED IN: mitochondrion, chloroplast; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: 6-phosphogluconate dehydrogenase, NAD-binding (InterPro:IPR006115), 6-phosphogluconate dehydrogenase, C-terminal-like (InterPro:IPR008927), Dehydrogenase, multihelical (InterPro:IPR013328), 6-phosphogluconate dehydrogenase, decarboxylating (InterPro:IPR006113), 6-phosphogluconate dehydrogenase, C-terminal (InterPro:IPR006114), 6-phosphogluconate dehydrogenase (InterPro:IPR006183), NAD(P) binding (InterPro:IPR016040), Fibritin/6-phosphogluconate dehydrogenase, C-terminal extension (InterPro:IPR012284); BEST Arabidopsis thaliana protein match is: 6-phosphogluconate dehydrogenase family protein (TAIR:AT1G64190.1); Has 8481 Blast hits to 8387 proteins in 1495 species: Archae - 46; Bacteria - 4731; Metazoa - 542; Fungi - 166; Plants - 211; Viruses - 2; Other |

| | | | | | |
|---|---|---|---|---|---|
| Gma.10820.2.S1_at | 0.03369 | -1.6177 | Glyma17g34690.1 | AT2G26080.1 | Arabidopsis thaliana glycine decarboxylase P-protein 2 (AtGLDP2); FUNCTIONS IN: glycine dehydrogenase (decarboxylating) activity, ATP binding; INVOLVED IN: glycine decarboxylation via glycine cleavage system, glycine metabolic process; LOCATED IN: mitochondrion, glycine cleavage complex, chloroplast, chloroplast envelope; EXPRESSED IN: 33 plant structures; EXPRESSED DURING: 16 growth stages; CONTAINS InterPro DOMAIN/s: Pyridoxal phosphate-dependent transferase, major region (InterPro:IPR015424), Glycine cleavage system P-protein (InterPro:IPR003437), Pyridoxal phosphate-dependent transferase, major region, subdomain 1 (InterPro:IPR015421); BEST Arabidopsis thaliana protein match is: AtGLDP1 (Arabidopsis thaliana glycine decarboxylase P-protein 1); catalytic/ glycine dehydrogenase (decarboxylating)/ pyridoxal phosphate binding (TAIR:AT4G33010.1); Has 9803 Blast hits to 8963 proteins in 1111 species: Archae - 132; Bacteria - 2858; Metazoa - 123; Fungi - 158; Plants - 73; Viruses - 0; Other Eukaryotes - 6459 (source: NCBI BLink). |
| GmaAffx.83343.1.S1_ | 0.06717 | -1.615 | Glyma11g15010.1 | AT2G28760.3 | UDP-XYL SYNTHASE 6 (UXS6); FUNCTIONS IN: coenzyme binding, binding, catalytic activity; INVOLVED IN: cellular metabolic process, nucleotide-sugar metabolic process, metabolic process; LOCATED IN: plasma membrane; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: NAD-dependent epimerase/dehydratase (InterPro:IPR001509), NAD(P)-binding (InterPro:IPR016040); BEST Arabidopsis thaliana protein match is: UXS5; UDP-glucuronate decarboxylase/ catalytic (TAIR:AT3G46440.2); Has 32193 Blast hits to 32159 proteins in 1749 species: Archae - 518; Bacteria - 14743; Metazoa - 694; Fungi - 281; Plants - 822; Viruses - 72; Other Eukaryotes - 15063 (source: NCBI BLink). |
| Gma.618.1.S1_at | 0.05225 | -1.6124 | Glyma09g21050.1 | AT2G01350.1 | At2g01350 encodes quinolinate phosphoribosyl transferase involved in NAD biosynthesis as shown by heterologous expression in E. coli. |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.41477.1.A1 | 0.06449 | -1.6031 | Glyma17g14760.1 | AT1G12230.1 | transaldolase, putative; FUNCTIONS IN: transaldolase activity, catalytic activity; INVOLVED IN: glucose catabolic process to lactate and acetate, 5-phosphoribose 1-diphosphate biosynthetic process, carbohydrate metabolic process, metabolic process, pentose-phosphate shunt, non-oxidative branch; LOCATED IN: chloroplast; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Aldolase-type TIM barrel (InterPro:IPR013785), Transaldolase (InterPro:IPR001585); Has 4687 Blast hits to 4687 proteins in 1248 species: Archae - 40; Bacteria - 2981; Metazoa - 141; Fungi - 151; Plants - 60; Viruses - 7; Other Eukaryotes - 1307 (source: NCBI BLink). |
| Gma.5859.1.S1_at | 0.08446 | -1.6014 | Glyma08g13290.1 | AT2G31810.1 | acetolactate synthase small subunit, putative; FUNCTIONS IN: acetolactate synthase activity, amino acid binding; INVOLVED IN: branched chain family amino acid biosynthetic process; LOCATED IN: chloroplast; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Acetolactate synthase, small subunit (InterPro:IPR004789), Amino acid-binding ACT (InterPro:IPR002912); BEST Arabidopsis thaliana protein match is: acetolactate synthase small subunit, putative (TAIR:AT5G16290.2); Has 8371 Blast hits to 4290 proteins in 1102 species: Archae - 146; Bacteria - 3993; Metazoa - 0; Fungi - 172; Plants - 54; Viruses - 0; Other Eukaryotes - 4006 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.1394.1.S1_at | 0.07258 | -1.6011 | Glyma16g02110.2 | AT1G32900.1 | starch synthase, putative; FUNCTIONS IN: transferase activity, transferring glycosyl groups; INVOLVED IN: biosynthetic process, glucan biosynthetic process; LOCATED IN: chloroplast; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Glycogen/starch synthases, ADP-glucose type (InterPro:IPR011835), Starch synthase catalytic region (InterPro:IPR013534), Glycosyl transferase, group 1 (InterPro:IPR001296); BEST Arabidopsis thaliana protein match is: AtSS2 (starch synthase 2); transferase, transferring glycosyl groups (TAIR:AT3G01180.1); Has 9548 Blast hits to 9532 proteins in 2484 species: Archae - 212; Bacteria - 3723; Metazoa - 12; Fungi - 137; Plants - 4346; Viruses - 0; Other Eukaryotes - 1118 (source: NCBI BLink). |
| Gma.17932.1.A1_at | 0.05004 | -1.5912 | Glyma10g44180.2 | AT1G74040.1 | Encodes an active Arabidopsis isopropylmalate synthase IPMS2. Involved in leucine biosynthesis. Do not participate in the chain elongation of glucosinolates. Expressed constitutively throughout the plant. Loss of IPMS2 can be compensated by a second isopropylmalate synthase gene IPMS1 (At1g18500). |
| Gma.5248.1.S1_at | 0.09403 | -1.591 | Glyma13g44490.1 | AT1G80560.1 | The AtIMD2 is one out of 3 genes encoding the enzyme 3-isopropylmalate dehydrogenase involved in leucine biosynthesis in Arabidopsis. Its subcellular location has been targeted to plastids. |
| GmaAffx.84342.2.S1_ | 0.05654 | -1.5888 | Glyma08g04460.1 | AT4G13360.1 | catalytic; FUNCTIONS IN: catalytic activity; INVOLVED IN: metabolic process; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Crotonase, core (InterPro:IPR001753); BEST Arabidopsis thaliana protein match is: enoyl-CoA hydratase/isomerase family protein (TAIR:AT3G24360.1); Has 13706 Blast hits to 13704 proteins in 1059 species: Archae - 151; Bacteria - 8301; Metazoa - 722; Fungi - 359; Plants - 270; Viruses - 0; Other Eukaryotes - 3903 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.584.2.S1_a_at | 0.06238 | -1.5855 | Glyma05g27940.1 | AT4G20520.1 | RNA binding / RNA-directed DNA polymerase; FUNCTIONS IN: RNA binding, RNA-directed DNA polymerase activity; INVOLVED IN: RNA-dependent DNA replication; LOCATED IN: cellular_component unknown; CONTAINS InterPro DOMAIN/s: RNA-directed DNA polymerase (reverse transcriptase) (InterPro:IPR000477), Protein of unknown function DUF1204 (InterPro:IPR009596), RNA-directed DNA polymerase (reverse transcriptase), related (InterPro:IPR015706); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT4G20700.1); Has 1015 Blast hits to 1006 proteins in 90 species: Archae - 0; Bacteria - 56; Metazoa - 101; Fungi - 4; Plants - 834; Viruses - 0; Other Eukaryotes - 20 (source: NCBI BLink). |
| Gma.10730.1.S1_at | 0.07712 | -1.5846 | Glyma08g47990.1 | AT1G68590.1 | plastid-specific 30S ribosomal protein 3, putative / PSRP-3, putative; FUNCTIONS IN: structural constituent of ribosome; INVOLVED IN: translation; LOCATED IN: thylakoid, chloroplast thylakoid membrane, chloroplast; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Ribosomal protein, PSRP-3/Ycf65 (InterPro:IPR006924); BEST Arabidopsis thaliana protein match is: plastid-specific 30S ribosomal protein 3, putative / PSRP-3, putative (TAIR:AT5G15760.1); Has 334 Blast hits to 334 proteins in 87 species: Archae - 0; Bacteria - 107; Metazoa - 0; Fungi - 0; Plants - 49; Viruses - 0; Other Eukaryotes - 178 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.71056.1.S1_ | 0.09671 | -1.5743 | Ambiguous Hit | AT3G49100.1 | signal recognition particle 9 kDa protein, putative / SRP9, putative; FUNCTIONS IN: protein binding, 7S RNA binding, signal recognition particle binding; INVOLVED IN: negative regulation of translational elongation, SRP-dependent cotranslational protein targeting to membrane; LOCATED IN: signal recognition particle; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Signal recognition particle, SRP9 subunit (InterPro:IPR008832), Signal recognition particle, SRP9/SRP14 subunit (InterPro:IPR009018); Has 176 Blast hits to 176 proteins in 70 species: Archae - 0; Bacteria - 0; Metazoa - 123; Fungi - 10; Plants - 27; Viruses - 0; Other Eukaryotes - 16 (source: NCBI BLink). |
| Gma.325.1.S1_at | 0.07853 | -1.5742 | Glyma11g37250.1 | AT1G67550.1 | Encodes a nickel-containing urea hydrolase involved in nitrogen recycling. It requires three urease accessory proteins for its activation. |
| GmaAffx.91087.1.S1_ | 0.07716 | -1.5719 | No Soybean Match Identified, E<10E-30, Perecent ID>95% | AT4G11050.1 | Arabidopsis thaliana glycosyl hydrolase 9C3 (AtGH9C3); FUNCTIONS IN: carbohydrate binding, hydrolase activity, hydrolyzing O-glycosyl compounds, catalytic activity; INVOLVED IN: carbohydrate metabolic process; LOCATED IN: endomembrane system; EXPRESSED IN: 7 plant structures; EXPRESSED DURING: 4 anthesis, D bilateral stage, E expanded cotyledon stage; CONTAINS InterPro DOMAIN/s: Six-hairpin glycosidase (InterPro:IPR012341), Carbohydrate-binding (InterPro:IPR008965), Glycoside hydrolase, family 9, active site (InterPro:IPR018221), Six-hairpin glycosidase-like (InterPro:IPR008928), Glycoside hydrolase, family 9 (InterPro:IPR001701); BEST Arabidopsis thaliana protein match is: AtGH9C2 (Arabidopsis thaliana glycosyl hydrolase 9C2); carbohydrate binding / catalytic/ hydrolase, hydrolyzing O-glycosyl compounds (TAIR:AT1G64390.1); Has 1207 Blast hits to 1198 proteins in 196 species: Archae - 0; Bacteria - 379; Metazoa - 132; Fungi - 14; Plants - 625; Viruses - 0; Other Eukaryotes - 57 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.2876.1.S1_s_at | 0.07681 | -1.5708 | Glyma02g00810.1 | AT5G43330.1 | malate dehydrogenase, cytosolic, putative; FUNCTIONS IN: in 6 functions; INVOLVED IN: cellular carbohydrate metabolic process, glycolysis, malate metabolic process, carbohydrate metabolic process, metabolic process; LOCATED IN: plasma membrane; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Malate dehydrogenase, NAD-dependent, cytosolic (InterPro:IPR011274), Lactate/malate dehydrogenase (InterPro:IPR001236), Malate dehydrogenase, NAD or NADP (InterPro:IPR010945), NAD(P)-binding (InterPro:IPR016040), L-lactate/malate dehydrogenase (InterPro:IPR001557), Malate dehydrogenase, active site (InterPro:IPR001252), Lactate dehydrogenase/glycoside hydrolase, family 4, C-terminal (InterPro:IPR015955); BEST Arabidopsis thaliana protein match is: malate dehydrogenase, cytosolic, putative (TAIR:AT1G04410.1); Has 8255 Blast hits to 8253 proteins in 1762 species: Archae - 116; Bacteria - 4254; Metazoa - 1034; Fungi - 184; Plants - 466; Viruses - 0; Other Eukaryotes - 2201 (source: NCBI BLink). |
| Gma.17995.1.S1_at | 0.04759 | -1.5567 | Glyma05g03100.1 | AT5G03860.1 | Encodes a protein with malate synthase activity. |
| Gma.5728.1.S1_at | 0.09471 | -1.5399 | Glyma16g06700.3 | AT5G40810.1 | cytochrome c1, putative; FUNCTIONS IN: electron carrier activity, iron ion binding, heme binding, electron transporter, transferring electrons within CoQH2-cytochrome c reductase complex activity; LOCATED IN: mitochondrial respiratory chain, mitochondrion, vacuole, mitochondrial respiratory chain complex III, membrane; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Cytochrome c1 (InterPro:IPR002326), Cytochrome c, monohaem (InterPro:IPR009056); BEST Arabidopsis thaliana protein match is: cytochrome c1, putative (TAIR:AT3G27240.1); Has 2945 Blast hits to 2945 proteins in 515 species: Archae - 0; Bacteria - 722; Metazoa - 171; Fungi - 166; Plants - 63; Viruses - 0; Other Eukaryotes - 1823 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.13318.1.S1_at | 0.09268 | -1.5279 | Glyma15g07240.1 | AT4G05090.1 | inositol monophosphatase family protein; FUNCTIONS IN: 3'(2'),5'-bisphosphate nucleotidase activity, inositol or phosphatidylinositol phosphatase activity; INVOLVED IN: sulfur metabolic process; LOCATED IN: chloroplast; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Inositol monophosphatase (InterPro:IPR000760); BEST Arabidopsis thaliana protein match is: SAL1; 3'(2'),5'-bisphosphate nucleotidase/ inositol or phosphatidylinositol phosphatase (TAIR:AT5G63980.1); Has 3732 Blast hits to 3732 proteins in 818 species: Archae - 22; Bacteria - 1713; Metazoa - 80; Fungi - 151; Plants - 134; Viruses - 0; Other Eukaryotes - 1632 (source: NCBI BLink). |
| GmaAffx.56183.1.A1 | 0.04032 | -1.515 | Glyma10g04570.1 | AT5G44000.1 | glutathione S-transferase C-terminal domain-containing protein; FUNCTIONS IN: glutathione transferase activity; LOCATED IN: chloroplast; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Glutathione S-transferase, C-terminal (InterPro:IPR004046), Glutathione S-transferase, C-terminal-like (InterPro:IPR010987), Glutathione S-transferase/chloride channel, C-terminal (InterPro:IPR017933), Thioredoxin-like fold (InterPro:IPR012336); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT4G19880.1); Has 1652 Blast hits to 1652 proteins in 488 species: Archae - 12; Bacteria - 905; Metazoa - 22; Fungi - 164; Plants - 61; Viruses - 0; Other Eukaryotes - 488 (source: NCBI BLink). |
| Gma.5851.1.S1_at | 0.09399 | -1.5149 | Glyma11g10480.1 | AT2G16600.1 | Encodes cytosolic cyclophilin ROC3. |
| Gma.4803.1.A1_at | 0.05006 | -1.5129 | Glyma03g03210.1 | AT2G45280.1 | Encodes a protein similar to RAD51C involved in double stranded break repair via homologous recombination. Sensitive to DSB induced by Mitomycin C and gamma irradiation, interacts with Atxrcc3 in yeast two-hybrid assay. Required for female meiosis but not critical for mitosis under normal conditions. |

| | | | | | |
|---|---|---|---|---|---|
| Gma.2119.1.S1_at | 0.08382 | -1.5081 | No Soybean Match Identified, E<10E-30, Perecent ID>95% | AT5G03650.1 | Encodes starch branching enzyme (E.C.2.4.1.18) similar to SBE2 from maize and rice. Expressed throughout the plant and highest in seedlings and cauline leaves. |
| GmaAffx.91530.1.S1_ | 0.07443 | -1.5068 | Glyma01g36890.1 | AT5G66120.2 | 3-dehydroquinate synthase, putative; FUNCTIONS IN: 3-dehydroquinate synthase activity; INVOLVED IN: aromatic amino acid family biosynthetic process; LOCATED IN: chloroplast; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: 3-dehydroquinate synthase AroB, subgroup (InterPro:IPR016037), 3-dehydroquinate synthase AroB (InterPro:IPR002658); Has 5882 Blast hits to 5880 proteins in 1401 species: Archae - 124; Bacteria - 2775; Metazoa - 7; Fungi - 133; Plants - 27; Viruses - 2; Other Eukaryotes - 2814 (source: NCBI BLink). |
| Gma.11260.1.S1_x_at | 0.05226 | -1.5012 | Glyma03g37080.1 | AT3G12290.1 | tetrahydrofolate dehydrogenase/cyclohydrolase, putative; FUNCTIONS IN: binding, catalytic activity; INVOLVED IN: folic acid and derivative biosynthetic process, metabolic process; LOCATED IN: chloroplast; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Tetrahydrofolate dehydrogenase/cyclohydrolase (InterPro:IPR000672), NAD(P)-binding (InterPro:IPR016040); BEST Arabidopsis thaliana protein match is: tetrahydrofolate dehydrogenase/cyclohydrolase, putative (TAIR:AT4G00620.1); Has 7113 Blast hits to 7108 proteins in 1554 species: Archae - 77; Bacteria - 3117; Metazoa - 345; Fungi - 200; Plants - 87; Viruses - 0; Other Eukaryotes - 3287 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.12685.1.A1_at | 0.08224 | -1.5006 | Glyma04g03280.1 | AT1G75330.1 | ORNITHINE CARBAMOYLTRANSFERASE (OTC); FUNCTIONS IN: amino acid binding, ornithine carbamoyltransferase activity, carboxyl- or carbamoyltransferase activity; INVOLVED IN: amino acid metabolic process; LOCATED IN: chloroplast, chloroplast stroma; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Aspartate/ornithine carbamoyltransferase, carbamoyl-P binding (InterPro:IPR006132), Aspartate/ornithine carbamoyltransferase (InterPro:IPR006130); Aspartate/ornithine carbamoyltransferase, Asp/Orn-binding region (InterPro:IPR006131), Ornithine carbamoyltransferase (InterPro:IPR002292); BEST Arabidopsis thaliana protein match is: aspartate carabmoyltransferase, chloroplast / aspartate transcarbamylase / ATCase (PYRB) (TAIR:AT3G20330.1); Has 11337 Blast hits to 11337 proteins in 1659 species: Archae - 346; Bacteria - 6031; Metazoa - 180; Fungi - 195; Plants - 66; Viruses - 6; Other Eukaryotes - 4513 (source: NCBI BLink). |
| Gma.15061.1.S1_at | 0.08512 | -1.4982 | Glyma05g38510.1 | AT5G50920.1 | Encodes a protein that is similar to ATP-dependent Clp protease ATP-binding subunit / ClpC. Involved in protein import into the chloroplast. May provide ATP source that drives the TIC (Translocon at the Inner envelope membrane of Chloroplasts) translocation machinery. |

| | | | | | |
|---|---|---|---|---|---|
| Gma.11260.2.S1_at | 0.09985 | -1.4923 | Glyma19g39710.1 | AT3G12290.1 | tetrahydrofolate dehydrogenase/cyclohydrolase, putative; FUNCTIONS IN: binding, catalytic activity; INVOLVED IN: folic acid and derivative biosynthetic process, metabolic process; LOCATED IN: chloroplast; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Tetrahydrofolate dehydrogenase/cyclohydrolase (InterPro:IPR000672), NAD(P)-binding (InterPro:IPR016040); BEST Arabidopsis thaliana protein match is: tetrahydrofolate dehydrogenase/cyclohydrolase, putative (TAIR:AT4G00620.1); Has 7113 Blast hits to 7108 proteins in 1554 species: Archae - 77; Bacteria - 3117; Metazoa - 345; Fungi - 200; Plants - 87; Viruses - 0; Other Eukaryotes - 3287 (source: NCBI BLink). |
| Gma.9622.1.S1_s_at | 0.06995 | -1.4874 | Glyma04g40580.1 | AT5G54160.1 | A caffeic acid/5-hydroxyferulic acid O-methyltransferase. Interacts with 14-4-3 proteins in yeast 2 hybrid assay. AtOMT1 (At5g54160) encodes a flavonol 3'-O-methyltransferase that is highly active towards quercetin and myricetin. The substrate specificity identifies the enzyme as flavonol 3'-methyltransferase which replaces the former annotation of the gene to encode a caffeic acid/5-hydroxyferulic acid O-methyltransferase |
| Gma.4693.1.S1_at | 0.02679 | -1.4869 | Glyma14g35690.1 | AT3G21110.2 | 5'-phosphoribosyl-4-(N-succinocarboxamide)-5-aminoimidazole synthetase (PUR7, syn. SAICAR synthetase), catalyzes aspartate addition at the alpha-amino group to the growing purine backbone. |

| | | | | | |
|---|---|---|---|---|---|
| Gma.15810.1.A1_at | 0.08569 | -1.4865 | Glyma13g33070.1 | AT5G64360.4 | DNAJ heat shock N-terminal domain-containing protein; FUNCTIONS IN: heat shock protein binding; INVOLVED IN: protein folding; LOCATED IN: cellular_component unknown; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Molecular chaperone, heat shock protein, Hsp40, DnaJ (InterPro:IPR015609), Heat shock protein DnaJ, N-terminal (InterPro:IPR001623); BEST Arabidopsis thaliana protein match is: DNAJ heat shock N-terminal domain-containing protein (TAIR:AT5G09540.1); Has 1613 Blast hits to 1606 proteins in 464 species: Archae - 35; Bacteria - 630; Metazoa - 430; Fungi - 74; Plants - 291; Viruses - 0; Other Eukaryotes - 153 (source: NCBI BLink). |
| GmaAffx.63454.1.S1_ | 0.09268 | -1.4757 | Glyma13g35240.2 | AT5G59440.3 | Encodes thymidylate kinase which exists in two isoforms in plants. The longer variant of 263 amino acids with a N-terminal extension that is required for localization to the mitochondrion. The second isoform of 224 residues is localized to the cytoplasm and nucleoplasm. Peak of expression occurs during G1/S phase transition. |
| GmaAffx.46169.1.S1_ | 0.09248 | -1.4693 | Glyma07g06950.2 | AT3G61530.2 | Encodes a ketopentoate hydroxymethyltransferase that appears to localize to the mitochondria. This protein is expected to play a role in pantothenate (vitamin B5) biosynthesis. |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.41255.1.A1 | 0.09165 | -1.4687 | Glyma11g34400.1 | AT5G06060.1 | tropinone reductase, putative / tropine dehydrogenase, putative; FUNCTIONS IN: oxidoreductase activity, binding, catalytic activity; INVOLVED IN: metabolic process; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: NAD(P)-binding (InterPro:IPR016040), Glucose/ribitol dehydrogenase (InterPro:IPR002347), Short-chain dehydrogenase/reductase SDR (InterPro:IPR002198); BEST Arabidopsis thaliana protein match is: tropinone reductase, putative / tropine dehydrogenase, putative (TAIR:AT2G29290.2); Has 88350 Blast hits to 88156 proteins in 2259 species: Archae - 471; Bacteria - 47428; Metazoa - 5087; Fungi - 4564; Plants - 1650; Viruses - 7; Other Eukaryotes - 29143 (source: NCBI BLink). |
| Gma.2369.1.S1_at | 0.03272 | -1.4603 | Glyma05g33330.1 | AT5G61790.1 | calnexin 1 (CNX1); FUNCTIONS IN: unfolded protein binding, calcium ion binding; INVOLVED IN: protein folding; LOCATED IN: in 8 components; EXPRESSED IN: 26 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Calreticulin/calnexin, P (InterPro:IPR009033), Calreticulin/calnexin (InterPro:IPR001580), Calreticulin/calnexin, conserved site (InterPro:IPR018124), Concanavalin A-like lectin/glucanase (InterPro:IPR008985), Concanavalin A-like lectin/glucanase, subgroup (InterPro:IPR013320); BEST Arabidopsis thaliana protein match is: calnexin, putative (TAIR:AT5G07340.1); Has 1344 Blast hits to 1252 proteins in 297 species: Archae - 2; Bacteria - 61; Metazoa - 623; Fungi - 137; Plants - 191; Viruses - 36; Other Eukaryotes - 294 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.89539.1.S1_ | 0.03678 | -1.4584 | Glyma15g36190.3 | AT5G45775.2 | 60S ribosomal protein L11 (RPL11D); FUNCTIONS IN: structural constituent of ribosome; INVOLVED IN: translation; LOCATED IN: cytosolic large ribosomal subunit, vacuole; EXPRESSED IN: male gametophyte, leaf; CONTAINS InterPro DOMAIN/s: Ribosomal protein L5 (InterPro:IPR002132); BEST Arabidopsis thaliana protein match is: RPL16B; structural constituent of ribosome (TAIR:AT4G18730.1); Has 5263 Blast hits to 5263 proteins in 1657 species: Archae - 218; Bacteria - 2825; Metazoa - 209; Fungi - 107; Plants - 188; Viruses - 0; Other Eukaryotes - 1716 (source: NCBI BLink). |
| Gma.12657.2.A1_at | 0.0863 | -1.4512 | Glyma16g26340.1 | AT3G26710.1 | COFACTOR ASSEMBLY OF COMPLEX C (CCB1); FUNCTIONS IN: molecular_function unknown; INVOLVED IN: cytochrome b6f complex assembly; LOCATED IN: chloroplast; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; Has 168 Blast hits to 168 proteins in 43 species: Archae - 0; Bacteria - 60; Metazoa - 0; Fungi - 0; Plants - 23; Viruses - 0; Other Eukaryotes - 85 (source: NCBI BLink). |
| Gma.11052.1.S1_at | 0.0863 | -1.4376 | Glyma10g00790.1 | AT3G23490.1 | cyanase |
| GmaAffx.61298.1.S1_ | 0.03172 | -1.4182 | Glyma17g08010.1 | AT4G30690.1 | translation initiation factor 3 (IF-3) family protein; FUNCTIONS IN: translation initiation factor activity; INVOLVED IN: translational initiation; LOCATED IN: cell wall, chloroplast; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Initiation factor 3 (InterPro:IPR001288); BEST Arabidopsis thaliana protein match is: translation initiation factor 3 (IF-3) family protein (TAIR:AT2G24060.1); Has 5057 Blast hits to 5057 proteins in 1387 species: Archae - 0; Bacteria - 2592; Metazoa - 64; Fungi - 26; Plants - 43; Viruses - 0; Other Eukaryotes - 2332 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.69812.1.A1 | 0.05952 | -1.4169 | Glyma07g08980.1 | AT1G12550.1 | oxidoreductase family protein; FUNCTIONS IN: in 6 functions; INVOLVED IN: metabolic process; LOCATED IN: cellular_component unknown; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: D-isomer specific 2-hydroxyacid dehydrogenase, catalytic region (InterPro:IPR006139), D-isomer specific 2-hydroxyacid dehydrogenase, NAD-binding (InterPro:IPR006140), NAD(P)-binding (InterPro:IPR016040); BEST Arabidopsis thaliana protein match is: oxidoreductase family protein (TAIR:AT2G45630.2); Has 19677 Blast hits to 19673 proteins in 1516 species: Archae - 283; Bacteria - 9440; Metazoa - 658; Fungi - 751; Plants - 326; Viruses - 5; Other Eukaryotes - 8214 (source: NCBI BLink). |
| GmaAffx.41858.2.S1_ | 0.04862 | -1.4045 | Glyma08g03740.1 | AT4G29120.1 | 6-phosphogluconate dehydrogenase NAD-binding domain-containing protein; FUNCTIONS IN: coenzyme binding, oxidoreductase activity, phosphogluconate dehydrogenase (decarboxylating) activity, binding, catalytic activity; INVOLVED IN: pentose-phosphate shunt, metabolic process; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: 6-phosphogluconate dehydrogenase, NAD-binding (InterPro:IPR006115), 6-phosphogluconate dehydrogenase, C-terminal-like (InterPro:IPR008927), Dehydrogenase, multihelical (InterPro:IPR013328), 3-hydroxyacid dehydrogenase/reductase (InterPro:IPR015815), NAD(P)-binding (InterPro:IPR016040); BEST Arabidopsis thaliana protein match is: 6-phosphogluconate dehydrogenase NAD-binding domain-containing protein (TAIR:AT1G71180.1); Has 12459 Blast hits to 12442 proteins in 1308 species: Archae - 93; Bacteria - 6200; Metazoa - 393; Fungi - 342; Plants - 184; Viruses - 2; Other Eukaryotes - 5245 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.15879.1.A1_at | 0.09461 | -1.4033 | Glyma04g23550.1 | AT2G24060.1 | translation initiation factor 3 (IF-3) family protein; FUNCTIONS IN: translation initiation factor activity; INVOLVED IN: translational initiation; LOCATED IN: chloroplast; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Initiation factor 3 (InterPro:IPR001288); BEST Arabidopsis thaliana protein match is: translation initiation factor 3 (IF-3) family protein (TAIR:AT4G30690.1); Has 17514 Blast hits to 8661 proteins in 1545 species: Archae - 56; Bacteria - 6187; Metazoa - 2053; Fungi - 689; Plants - 758; Viruses - 336; Other Eukaryotes - 7435 (source: NCBI BLink). |
| GmaAffx.41858.1.S1_ | 0.07871 | -1.4032 | Glyma05g35880.1 | AT4G29120.1 | 6-phosphogluconate dehydrogenase NAD-binding domain-containing protein; FUNCTIONS IN: coenzyme binding, oxidoreductase activity, phosphogluconate dehydrogenase (decarboxylating) activity, binding, catalytic activity; INVOLVED IN: pentose-phosphate shunt, metabolic process; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: 6-phosphogluconate dehydrogenase, NAD-binding (InterPro:IPR006115), 6-phosphogluconate dehydrogenase, C-terminal-like (InterPro:IPR008927), Dehydrogenase, multihelical (InterPro:IPR013328), 3-hydroxyacid dehydrogenase/reductase (InterPro:IPR015815), NAD(P)-binding (InterPro:IPR016040); BEST Arabidopsis thaliana protein match is: 6-phosphogluconate dehydrogenase NAD-binding domain-containing protein (TAIR:AT1G71180.1); Has 12459 Blast hits to 12442 proteins in 1308 species: Archae - 93; Bacteria - 6200; Metazoa - 393; Fungi - 342; Plants - 184; Viruses - 2; Other Eukaryotes - 5245 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.17640.1.A1_at | 0.08139 | -1.3964 | Glyma08g20080.1 | AT1G22940.1 | Encodes a bifunctional enzyme required for thiamine (vitamin B1) biosynthesis. TH1 can phosphorylate HMP-P to produce HMP-PP, the pyrimidine heterocyclic subunit of thiamine. TH1 also catalyzes the condensation of HMP-PP and HET to form thiamine monophosphate (TMP). TH1 also appears capable of phosphorylating HMP based on E.coli mutant complementation assays. th1 mutants are thiamine auxotrophs that die as seedlings on unsupplemented media. |
| Gma.11330.2.S1_at | 0.09622 | -1.3918 | Glyma15g10210.1 | AT3G52580.1 | 40S ribosomal protein S14 (RPS14C); FUNCTIONS IN: structural constituent of ribosome; INVOLVED IN: translation; LOCATED IN: cytosolic small ribosomal subunit; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Ribosomal protein S11 (InterPro:IPR001971), Ribosomal S11, conserved site (InterPro:IPR018102); BEST Arabidopsis thaliana protein match is: 40S ribosomal protein S14 (RPS14B) (TAIR:AT3G11510.1); Has 6258 Blast hits to 6258 proteins in 1752 species: Archae - 169; Bacteria - 2837; Metazoa - 489; Fungi - 109; Plants - 512; Viruses - 0; Other Eukaryotes - 2142 (source: NCBI BLink). |
| GmaAffx.1301.127.A | 0.05376 | -1.3871 | Glyma04g43540.3 | AT2G44160.1 | methylenetetrahydrofolate reductase MTHFR2 mRNA, complete |
| GmaAffx.90866.1.S1_ | 0.09965 | -1.3805 | Glyma12g02790.1 | AT2G16600.1 | Encodes cytosolic cyclophilin ROC3. |
| Gma.3871.1.S1_at | 0.04164 | -1.3765 | Glyma07g10170.3 | AT3G18410.2 | NADH-ubiquinone oxidoreductase-related; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: photorespiration; LOCATED IN: mitochondrion, mitochondrial membrane, respiratory chain complex I; BEST Arabidopsis thaliana protein match is: NADH-ubiquinone oxidoreductase-related (TAIR:AT1G49140.1); Has 96 Blast hits to 96 proteins in 44 species: Archae - 0; Bacteria - 0; Metazoa - 6; Fungi - 54; Plants - 36; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.36115.1.A1 | 0.08151 | -1.3723 | Glyma04g40090.1 | AT2G40660.1 | tRNA-binding region domain-containing protein; FUNCTIONS IN: tRNA binding; INVOLVED IN: tRNA aminoacylation for protein translation; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Nucleic acid-binding, OB-fold-like (InterPro:IPR016027), Nucleic acid-binding, OB-fold (InterPro:IPR012340), tRNA-binding region (InterPro:IPR002547); BEST Arabidopsis thaliana protein match is: methionine--tRNA ligase, putative / methionyl-tRNA synthetase, putative / MetRS, putative (TAIR:AT4G13780.1); Has 3620 Blast hits to 3608 proteins in 1164 species: Archae - 158; Bacteria - 2173; Metazoa - 404; Fungi - 148; Plants - 87; Viruses - 1; Other Eukaryotes - 649 (source: NCBI BLink). |
| Gma.7460.1.S1_at | 0.03943 | -1.3711 | Glyma13g28910.2 | AT1G30000.1 | glycoside hydrolase family 47 protein; FUNCTIONS IN: mannosyl-oligosaccharide 1,2-alpha-mannosidase activity, alpha-mannosidase activity, calcium ion binding; INVOLVED IN: protein amino acid N-linked glycosylation; LOCATED IN: Golgi apparatus; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Glycoside hydrolase, family 47 (InterPro:IPR001382); BEST Arabidopsis thaliana protein match is: mannosyl-oligosaccharide 1,2-alpha-mannosidase, putative (TAIR:AT1G51590.1); Has 1507 Blast hits to 1364 proteins in 137 species: Archae - 0; Bacteria - 8; Metazoa - 700; Fungi - 557; Plants - 86; Viruses - 0; Other Eukaryotes - 156 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.1605.1.S1_x_at | 0.08357 | -1.3698 | No Soybean Match Identified, E<10E-30, Perecent ID>95% | AT1G48900.1 | signal recognition particle 54 kDa protein 3 / SRP54 (SRP-54C); FUNCTIONS IN: 7S RNA binding, mRNA binding, nucleoside-triphosphatase activity, GTP binding, nucleotide binding; INVOLVED IN: SRP-dependent cotranslational protein targeting to membrane, signal sequence recognition, SRP-dependent cotranslational protein targeting to membrane; LOCATED IN: signal recognition particle, signal recognition particle, endoplasmic reticulum targeting; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: ATPase, AAA+ type, core (InterPro:IPR003593), Signal recognition particle, SRP54 subunit, helical bundle (InterPro:IPR013822), Signal recognition particle, SRP54 subunit, M-domain (InterPro:IPR004125), Signal recognition particle, SRP54 subunit (InterPro:IPR006325), Signal recognition particle, SRP54 subunit, GTPase (InterPro:IPR000897); BEST Arabidopsis thaliana protein match is: signal recognition particle 54 kDa protein 2 / SRP54 (SRP-54B) (TAIR:AT5G49500.1); Has 12023 Blast hits to 12019 proteins in 1590 species: Archae - 312; Bacteria - 5666; Metazoa - |
| Gma.10752.1.S1_at | 0.07128 | -1.3698 | Glyma11g08050.1 | AT4G36130.1 | 60S ribosomal protein L8 (RPL8C); FUNCTIONS IN: structural constituent of ribosome; INVOLVED IN: translation; LOCATED IN: cytosolic ribosome, cytosolic large ribosomal subunit, ribosome, plasma membrane, vacuole; EXPRESSED IN: guard cell, cultured cell; CONTAINS InterPro DOMAIN/s: Translation protein SH3-like (InterPro:IPR008991), Nucleic acid-binding, OB-fold-like (InterPro:IPR016027), Nucleic acid-binding, OB-fold (InterPro:IPR012340), Ribosomal protein L2, domain 3 (InterPro:IPR014726), Ribosomal protein L2 (InterPro:IPR002171); BEST Arabidopsis thaliana protein match is: EMB2296 (embryo defective 2296); structural constituent of ribosome (TAIR:AT2G18020.1); Has 7437 Blast hits to 7435 proteins in 2204 species: Archae - 236; Bacteria - 3125; Metazoa - 339; Fungi - 188; Plants - 929; Viruses - 0; Other Eukaryotes - 2620 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.11295.2.S1_at | 0.05226 | -1.3674 | Glyma01g41810.1 | AT5G56670.1 | 40S ribosomal protein S30 (RPS30C); FUNCTIONS IN: structural constituent of ribosome; INVOLVED IN: translation; LOCATED IN: cytosolic small ribosomal subunit, ribosome; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Ribosomal protein S30 (InterPro:IPR006846); BEST Arabidopsis thaliana protein match is: 40S ribosomal protein S30 (RPS30B) (TAIR:AT4G29390.1); Has 487 Blast hits to 487 proteins in 186 species: Archae - 2; Bacteria - 0; Metazoa - 223; Fungi - 90; Plants - 56; Viruses - 1; Other Eukaryotes - 115 (source: NCBI BLink). |
| Gma.937.1.A1_at | 0.06494 | -1.3669 | No Soybean Match Identified, E<10E-30, Perecent ID>95% | AT5G62530.1 | Encodes mitochondrial Delta-pyrroline-5- carboxylate dehydrogenase. Involved in the catabolism of proline to glutamate. Involved in protection from proline toxicity. Induced at pathogen infection sites. P5CDH and SRO5 (an overlapping gene in the sense orientation) generate 24-nt and 21-nt siRNAs, which together are components of a regulatory loop controlling reactive oxygen species (ROS) production and stress response. |
| Gma.1563.1.S1_at | 0.05603 | -1.3647 | Glyma11g11710.1 | AT5G16650.1 | DNAJ heat shock N-terminal domain-containing protein; FUNCTIONS IN: unfolded protein binding, heat shock protein binding; INVOLVED IN: protein folding; LOCATED IN: cellular_component unknown; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Molecular chaperone, heat shock protein, Hsp40, DnaJ (InterPro:IPR015609), Heat shock protein DnaJ, N-terminal (InterPro:IPR001623), Heat shock protein DnaJ (InterPro:IPR003095); BEST Arabidopsis thaliana protein match is: DNAJ heat shock N-terminal domain-containing protein (TAIR:AT2G33735.1); Has 16083 Blast hits to 16080 proteins in 1936 species: Archae - 108; Bacteria - 5268; Metazoa - 3302; Fungi - 1476; Plants - 1210; Viruses - 11; Other Eukaryotes - 4708 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.11260.1.S1_at | 0.07072 | -1.3503 | Glyma03g37080.1 | AT3G12290.1 | tetrahydrofolate dehydrogenase/cyclohydrolase, putative; FUNCTIONS IN: binding, catalytic activity; INVOLVED IN: folic acid and derivative biosynthetic process, metabolic process; LOCATED IN: chloroplast; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Tetrahydrofolate dehydrogenase/cyclohydrolase (InterPro:IPR000672), NAD(P)-binding (InterPro:IPR016040); BEST Arabidopsis thaliana protein match is: tetrahydrofolate dehydrogenase/cyclohydrolase, putative (TAIR:AT4G00620.1); Has 7113 Blast hits to 7108 proteins in 1554 species: Archae - 77; Bacteria - 3117; Metazoa - 345; Fungi - 200; Plants - 87; Viruses - 0; Other Eukaryotes - 3287 (source: NCBI BLink). |
| Gma.2150.1.S1_at | 0.05654 | -1.3484 | Glyma11g07540.1 | AT1G19580.1 | Encodes mitochondrial gamma carbonic anhydrase. Component of the NADH dehydrogenase complex. |
| Gma.10820.2.S1_s_at | 0.06366 | -1.3441 | Glyma17g34690.1 | AT2G26080.1 | Arabidopsis thaliana glycine decarboxylase P-protein 2 (AtGLDP2); FUNCTIONS IN: glycine dehydrogenase (decarboxylating) activity, ATP binding; INVOLVED IN: glycine decarboxylation via glycine cleavage system, glycine metabolic process; LOCATED IN: mitochondrion, glycine cleavage complex, chloroplast, chloroplast envelope; EXPRESSED IN: 33 plant structures; EXPRESSED DURING: 16 growth stages; CONTAINS InterPro DOMAIN/s: Pyridoxal phosphate-dependent transferase, major region (InterPro:IPR015424), Glycine cleavage system P-protein (InterPro:IPR003437), Pyridoxal phosphate-dependent transferase, major region, subdomain 1 (InterPro:IPR015421); BEST Arabidopsis thaliana protein match is: AtGLDP1 (Arabidopsis thaliana glycine decarboxylase P-protein 1); catalytic/ glycine dehydrogenase (decarboxylating)/ pyridoxal phosphate binding (TAIR:AT4G33010.1); Has 9803 Blast hits to 8963 proteins in 1111 species: Archae - 132; Bacteria - 2858; Metazoa - 123; Fungi - 158; Plants - 73; Viruses - 0; Other Eukaryotes - 6459 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.11328.1.S1_at | 0.06847 | -1.3427 | Glyma12g05280.2 | AT2G30695.2 | FUNCTIONS IN: molecular_function unknown; INVOLVED IN: protein folding, protein transport; LOCATED IN: chloroplast; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Trigger factor, ribosome-binding, bacterial (InterPro:IPR008881); Has 207 Blast hits to 207 proteins in 55 species: Archae - 0; Bacteria - 110; Metazoa - 0; Fungi - 0; Plants - 22; Viruses - 0; Other Eukaryotes - 75 (source: NCBI BLink). |
| Gma.4256.1.S1_at | 0.07376 | -1.3388 | Glyma06g03340.1 | AT1G27680.1 | ADP-glucose pyrophosphorylase catalyzes the first, rate limiting step in starch biosynthesis. The large subunit plays a regulatory role whereas the small subunit (ApS) is the catalytic isoform. Four isoforms of the large subunit (ApL1-4) have been described.Mutational analysis of APS1 suggests that APL1 and APL2 can compensate for loss of APS1 catalytic activity,suggesting both have catalytic as well as regulatory functions. |
| Gma.17876.1.A1_at | 0.08139 | -1.3205 | Glyma19g43270.1 | AT3G04820.1 | pseudouridine synthase; FUNCTIONS IN: pseudouridine synthase activity; INVOLVED IN: tRNA processing, pseudouridine synthesis, tRNA pseudouridine synthesis; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: tRNA pseudouridine synthase D (InterPro:IPR001656), tRNA pseudouridine synthase D, eukaryotic (InterPro:IPR017091), tRNA pseudouridine synthase D, core (InterPro:IPR011760); Has 1538 Blast hits to 1478 proteins in 609 species: Archae - 166; Bacteria - 781; Metazoa - 175; Fungi - 102; Plants - 30; Viruses - 0; Other Eukaryotes - 284 (source: NCBI BLink). |
| GmaAffx.377.2.S1_at | 0.09903 | -1.3134 | Glyma03g39890.1 | AT5G27450.2 | Encodes a protein with mevalonate kinase activity involved in the mevalonate pathway. |

| | | | | | |
|---|---|---|---|---|---|
| Gma.15823.1.S1_at | 0.08755 | -1.3073 | No Soybean Match Identified, E<10E-30, Perecent ID>95% | AT3G59540.1 | 60S ribosomal protein L38 (RPL38B); FUNCTIONS IN: structural constituent of ribosome; INVOLVED IN: translation, ribosome biogenesis; LOCATED IN: cytosolic large ribosomal subunit, ribosome; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Ribosomal protein L38e (InterPro:IPR002675); BEST Arabidopsis thaliana protein match is: 60S ribosomal protein L38 (RPL38A) (TAIR:AT2G43460.1); Has 458 Blast hits to 458 proteins in 191 species: Archae - 3; Bacteria - 0; Metazoa - 220; Fungi - 79; Plants - 70; Viruses - 0; Other Eukaryotes - 86 (source: NCBI BLink). |
| Gma.6740.1.S1_x_at | 0.08623 | -1.2953 | Glyma08g19320.1 | AT4G14320.1 | 60S ribosomal protein L36a/L44 (RPL36aB); FUNCTIONS IN: structural constituent of ribosome; INVOLVED IN: translation; LOCATED IN: cytosolic large ribosomal subunit, ribosome; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Ribosomal protein L44e (InterPro:IPR000552), Ribosomal protein, zinc-binding (InterPro:IPR011332); BEST Arabidopsis thaliana protein match is: 60S ribosomal protein L36a/L44 (RPL36aA) (TAIR:AT3G23390.1); Has 750 Blast hits to 748 proteins in 259 species: Archae - 104; Bacteria - 1; Metazoa - 299; Fungi - 120; Plants - 71; Viruses - 0; Other Eukaryotes - 155 (source: NCBI BLink). |
| GmaAffx.84741.1.A1 | 0.08151 | -1.2943 | Glyma06g47860.1 | AT1G69485.1 | structural constituent of ribosome; FUNCTIONS IN: structural constituent of ribosome; INVOLVED IN: translation; LOCATED IN: large ribosomal subunit; CONTAINS InterPro DOMAIN/s: Ribosomal protein L32p (InterPro:IPR002677); BEST Arabidopsis thaliana protein match is: structural constituent of ribosome (TAIR:AT1G26740.1); Has 48 Blast hits to 48 proteins in 16 species: Archae - 0; Bacteria - 8; Metazoa - 0; Fungi - 2; Plants - 31; Viruses - 0; Other Eukaryotes - 7 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.1605.1.S1_s_at | 0.09622 | -1.2936 | No Soybean Match Identified, E<10E-30, Perecent ID>95% | AT1G48900.1 | signal recognition particle 54 kDa protein 3 / SRP54 (SRP-54C); FUNCTIONS IN: 7S RNA binding, mRNA binding, nucleoside-triphosphatase activity, GTP binding, nucleotide binding; INVOLVED IN: SRP-dependent cotranslational protein targeting to membrane, signal sequence recognition, SRP-dependent cotranslational protein targeting to membrane; LOCATED IN: signal recognition particle, signal recognition particle, endoplasmic reticulum targeting; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: ATPase, AAA+ type, core (InterPro:IPR003593), Signal recognition particle, SRP54 subunit, helical bundle (InterPro:IPR013822), Signal recognition particle, SRP54 subunit, M-domain (InterPro:IPR004125), Signal recognition particle, SRP54 subunit (InterPro:IPR006325), Signal recognition particle, SRP54 subunit, GTPase (InterPro:IPR000897); BEST Arabidopsis thaliana protein match is: signal recognition particle 54 kDa protein 2 / SRP54 (SRP-54B) (TAIR:AT5G49500.1); Has 12023 Blast hits to 12019 proteins in 1590 species: Archae - 312; Bacteria - 5666; Metazoa - |
| Gma.10596.1.S1_at | 0.08139 | -1.291 | Ambiguous Hit | AT3G59540.1 | 60S ribosomal protein L38 (RPL38B); FUNCTIONS IN: structural constituent of ribosome; INVOLVED IN: translation, ribosome biogenesis; LOCATED IN: cytosolic large ribosomal subunit, ribosome; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Ribosomal protein L38e (InterPro:IPR002675); BEST Arabidopsis thaliana protein match is: 60S ribosomal protein L38 (RPL38A) (TAIR:AT2G43460.1); Has 458 Blast hits to 458 proteins in 191 species: Archae - 3; Bacteria - 0; Metazoa - 220; Fungi - 79; Plants - 70; Viruses - 0; Other Eukaryotes - 86 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.45058.2.S1_ | 0.09471 | -1.2659 | Glyma13g41070.1 | AT2G45910.1 | protein kinase family protein / U-box domain-containing protein; FUNCTIONS IN: ubiquitin-protein ligase activity, protein serine/threonine kinase activity, protein kinase activity, kinase activity, ATP binding; INVOLVED IN: protein amino acid phosphorylation, protein ubiquitination; LOCATED IN: ubiquitin ligase complex; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Protein kinase, ATP binding site (InterPro:IPR017441), Protein kinase, core (InterPro:IPR000719), U box (InterPro:IPR003613), Serine/threonine protein kinase-related (InterPro:IPR017442), Protein kinase-like (InterPro:IPR011009), Serine/threonine protein kinase, active site (InterPro:IPR008271); BEST Arabidopsis thaliana protein match is: protein kinase family protein / U-box domain-containing protein (TAIR:AT3G49060.1); Has 92032 Blast hits to 90051 proteins in 3362 species: Archae - 123; Bacteria - 8431; Metazoa - 40368; Fungi - 7218; Plants - 19640; Viruses - 399; Other Eukaryotes - 15853 (source: NCBI BLink). |
| Gma.4375.2.S1_at | 0.09622 | -1.2412 | Glyma18g04660.1 | AT5G37600.1 | encodes a cytosolic glutamine synthetase, the enzyme has high affinity with substrate ammonium |
| Gma.9266.1.S1_s_at | 0.09965 | -1.2237 | Glyma04g02100.1 | AT5G42270.1 | VAR1 contains a conserved motif for ATPase and a metalloprotease characteristic to FtsH proteins, and is targeted into chloroplasts. A VAR1-fusion protein synthesized in vitro exhibited ATPase activity and partial metalloprotease activity. This protein is located to the thylakoid membrane and forms a complex with VAR2. FtsH1 (VAR1) and FtsH5 are interchangeable in thylakoid membranes. |

**Table 4 Other Candidate genes**

| Affy ID | Q-value | Fold change | Soybean Gene | Arabidopsis homolog | Description |
|---|---|---|---|---|---|
| Gma.2912.1.S1_at | 0.04466 | 26.819 | No Soybean Match Identified, E<10E-30, Perecent ID>95% | AT1G52690.2 | late embryogenesis abundant protein, putative / LEA protein, putative; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: embryonic development ending in seed dormancy; LOCATED IN: cellular_component unknown; EXPRESSED IN: 9 plant structures; EXPRESSED DURING: dry seed stage, 4 anthesis, LP.04 four leaves visible, petal differentiation and expansion stage; BEST Arabidopsis thaliana protein match is: late embryogenesis abundant protein, putative / LEA protein, putative (TAIR:AT3G15670.1); Has 2025 Blast hits to 947 proteins in 265 species: Archae - 6; Bacteria - 474; Metazoa - 312; Fungi - 113; Plants - 812; Viruses - 15; Other Eukaryotes - 293 (source: NCBI BLink). |
| Gma.2912.1.S1_s_at | 0.07148 | 11.377 | Glyma13g44020.1 | AT1G52690.2 | late embryogenesis abundant protein, putative / LEA protein, putative; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: embryonic development ending in seed dormancy; LOCATED IN: cellular_component unknown; EXPRESSED IN: 9 plant structures; EXPRESSED DURING: dry seed stage, 4 anthesis, LP.04 four leaves visible, petal differentiation and expansion stage; BEST Arabidopsis thaliana protein match is: late embryogenesis abundant protein, putative / LEA protein, putative (TAIR:AT3G15670.1); Has 2025 Blast hits to 947 proteins in 265 species: Archae - 6; Bacteria - 474; Metazoa - 312; Fungi - 113; Plants - 812; Viruses - 15; Other Eukaryotes - 293 (source: NCBI BLink). |
| GmaAffx.74591.1.S1_at | 0.00526 | 9.4551 | No Soybean Match Identified, E<10E-30, Perecent ID>95% | AT3G12490.1 | Encodes a protein with cysteine proteinase inhibitor activity. Overexpression increases tolerance to abiotic stressors (i.e.salt,osmitic, cold stress). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.8528.1.S1_at | 0.07445 | 8.2949 | Glyma20g28550.1 | AT3G22490.1 | late embryogenesis abundant protein, putative / LEA protein, putative; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: embryonic development ending in seed dormancy; LOCATED IN: cellular_component unknown; CONTAINS InterPro DOMAIN/s: Seed maturation protein (InterPro:IPR007011); BEST Arabidopsis thaliana protein match is: ATECP31 (TAIR:AT3G22500.1); Has 125 Blast hits to 114 proteins in 29 species: Archae - 2; Bacteria - 32; Metazoa - 7; Fungi - 0; Plants - 80; Viruses - 3; Other Eukaryotes - 1 (source: NCBI BLink). |
| Gma.13388.1.S1_s_at | 0.08139 | 7.3101 | Glyma14g03180.1 | AT2G42870.1 | Encodes PHYTOCHROME RAPIDLY REGULATED1 (PAR1), an atypical basic helix-loop-helix (bHLP) protein. Closely related to PAR2 (At3g58850). Up regulated after simulated shade perception. Acts in the nucleus to control plant development and as a negative regulator of shade avoidance response. Functions as transcriptional repressor of auxin-responsive genes SAUR15 (AT4G38850) and SAUR68 (AT1G29510). |
| GmaAffx.79054.1.S1_at | 0.05952 | 6.3055 | Glyma02g37060.1 | AT4G08950.1 | EXORDIUM (EXO); FUNCTIONS IN: molecular_function unknown; INVOLVED IN: response to brassinosteroid stimulus; LOCATED IN: cell wall, plant-type cell wall; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Phosphate-induced protein 1 conserved region (InterPro:IPR006766); BEST Arabidopsis thaliana protein match is: PHI-1 (PHOSPHATE-INDUCED 1) (TAIR:AT1G35140.1); Has 234 Blast hits to 233 proteins in 17 species: Archae - 0; Bacteria - 2; Metazoa - 0; Fungi - 0; Plants - 232; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.8411.1.S1_at | 0.02679 | 6.0034 | Glyma08g18890.1 | AT5G65390.1 | AGP7; LOCATED IN: anchored to membrane; EXPRESSED IN: 18 plant structures; EXPRESSED DURING: 12 growth stages; Has 67907 Blast hits to 26759 proteins in 1358 species: Archae - 408; Bacteria - 17588; Metazoa - 17399; Fungi - 6154; Plants - 9463; Viruses - 2368; Other Eukaryotes - 14527 (source: NCBI BLink). |
| Gma.772.1.S1_at | 0.03725 | 5.91 | No Soybean Match Identified, E<10E-30, Perecent ID>95% | AT5G12010.1 | unknown protein; INVOLVED IN: response to salt stress; LOCATED IN: chloroplast, plasma membrane, membrane; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT4G29780.1); Has 746 Blast hits to 744 proteins in 70 species: Archae - 0; Bacteria - 0; Metazoa - 425; Fungi - 40; Plants - 268; Viruses - 0; Other Eukaryotes - 13 (source: NCBI BLink). |
| Gma.9866.1.S1_at | 0.05226 | 5.6656 | Glyma11g10320.1 | AT1G76160.1 | SKU5 Similar 5 (sks5); FUNCTIONS IN: oxidoreductase activity, copper ion binding; LOCATED IN: apoplast, cell wall, plant-type cell wall; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Multicopper oxidase, type 3 (InterPro:IPR011707), Cupredoxin (InterPro:IPR008972), Multicopper oxidase, type 2 (InterPro:IPR011706), Multicopper oxidase, type 1 (InterPro:IPR001117); BEST Arabidopsis thaliana protein match is: SKS6 (SKU5-SIMILAR 6); pectinesterase (TAIR:AT1G41830.1); Has 3613 Blast hits to 3566 proteins in 612 species: Archae - 8; Bacteria - 911; Metazoa - 254; Fungi - 1504; Plants - 791; Viruses - 0; Other Eukaryotes - 145 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.6778.1.A1_at | 0.00857 | 5.5762 | Glyma06g02900.1 | AT1G75620.1 | glyoxal oxidase-related; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: endomembrane system; EXPRESSED IN: flower, root; EXPRESSED DURING: 4 anthesis; CONTAINS InterPro DOMAIN/s: Galactose oxidase/kelch, beta-propeller (InterPro:IPR011043), Galactose oxidase, beta-propeller (InterPro:IPR015916), Immunoglobulin E-set (InterPro:IPR014756), Glyoxal oxidase, N-terminal (InterPro:IPR009880), Region of unknown function DUF1929 (InterPro:IPR015202); BEST Arabidopsis thaliana protein match is: glyoxal oxidase-related (TAIR:AT1G19900.1); Has 563 Blast hits to 561 proteins in 114 species: Archae - 0; Bacteria - 228; Metazoa - 0; Fungi - 169; Plants - 152; Viruses - 0; Other Eukaryotes - 14 (source: NCBI BLink). |
| Gma.8859.1.A1_at | 0.03369 | 5.1687 | Glyma15g03110.1 | AT1G72790.1 | hydroxyproline-rich glycoprotein family protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cytosol, nucleus, plasma membrane; EXPRESSED IN: 14 plant structures; EXPRESSED DURING: 9 growth stages; BEST Arabidopsis thaliana protein match is: hydroxyproline-rich glycoprotein family protein (TAIR:AT5G57070.1); Has 15974 Blast hits to 9634 proteins in 554 species: Archae - 27; Bacteria - 1047; Metazoa - 6923; Fungi - 2241; Plants - 3444; Viruses - 523; Other Eukaryotes - 1769 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.2331.1.S1_at | 0.09106 | 5.0605 | Glyma06g09810.1 | AT5G49700.1 | DNA-binding protein-related; INVOLVED IN: biological_process unknown; EXPRESSED IN: shoot apex, embryo, inflorescence meristem; EXPRESSED DURING: D bilateral stage; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF296 (InterPro:IPR005175), Predicted AT-hook DNA-binding (InterPro:IPR014476); BEST Arabidopsis thaliana protein match is: DNA-binding protein-related (TAIR:AT1G14490.1); Has 424 Blast hits to 423 proteins in 20 species: Archae - 0; Bacteria - 2; Metazoa - 3; Fungi - 2; Plants - 417; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| Gma.4470.1.S1_s_at | 0.03377 | 5.0347 | Glyma01g26840.2 | AT3G54040.1 | photoassimilate-responsive protein-related; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: endomembrane system; EXPRESSED IN: 20 plant structures; EXPRESSED DURING: 10 growth stages; CONTAINS InterPro DOMAIN/s: PAR1 (InterPro:IPR009489); BEST Arabidopsis thaliana protein match is: photoassimilate-responsive protein, putative (TAIR:AT5G52390.1); Has 27 Blast hits to 27 proteins in 6 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 27; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.69994.1.S1_at | 0.08512 | 4.1721 | Glyma14g35330.1 | AT4G08950.1 | EXORDIUM (EXO); FUNCTIONS IN: molecular_function unknown; INVOLVED IN: response to brassinosteroid stimulus; LOCATED IN: cell wall, plant-type cell wall; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Phosphate-induced protein 1 conserved region (InterPro:IPR006766); BEST Arabidopsis thaliana protein match is: PHI-1 (PHOSPHATE-INDUCED 1) (TAIR:AT1G35140.1); Has 234 Blast hits to 233 proteins in 17 species: Archae - 0; Bacteria - 2; Metazoa - 0; Fungi - 0; Plants - 232; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| GmaAffx.71627.1.S1_at | 0.05811 | 4.0429 | Glyma03g39180.2 | AT5G57580.1 | calmodulin-binding protein; FUNCTIONS IN: calmodulin binding; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Calmodulin binding protein-like (InterPro:IPR012416); BEST Arabidopsis thaliana protein match is: calmodulin-binding protein (TAIR:AT4G25800.2); Has 187 Blast hits to 178 proteins in 13 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 187; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| GmaAffx.88340.1.S1_at | 0.04477 | 4.0429 | Glyma18g01220.1 | AT1G67570.1 | unknown protein; LOCATED IN: nucleus; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT1G50630.1); Has 81 Blast hits to 81 proteins in 8 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 79; Viruses - 0; Other Eukaryotes - 2 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.3594.2.S1_x_at | 0.03369 | 3.9901 | Glyma17g17970.1 | AT4G37450.1 | AGP18 is a lysine-rich arabinogalactan-protein (AGP) and part of a multi-gene family of glycoproteins with approx. 50 members. It falls into one subclass with AGP17 and AGP19, other lysine-rich AGPs. It is expressed in young leaves, shoots, roots and flowers. |
| Gma.3594.2.S1_at | 0.02679 | 3.8609 | Glyma17g17970.1 | AT4G37450.1 | AGP18 is a lysine-rich arabinogalactan-protein (AGP) and part of a multi-gene family of glycoproteins with approx. 50 members. It falls into one subclass with AGP17 and AGP19, other lysine-rich AGPs. It is expressed in young leaves, shoots, roots and flowers. |
| GmaAffx.66669.1.S1_at | 0.07712 | 3.8065 | Glyma20g35670.1 | AT1G33420.1 | PHD finger family protein; FUNCTIONS IN: protein binding, DNA binding, zinc ion binding; INVOLVED IN: regulation of transcription, DNA-dependent; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Zinc finger, RING/FYVE/PHD-type (InterPro:IPR013083), Zinc finger, PHD-type (InterPro:IPR001965), Zinc finger, FYVE/PHD-type (InterPro:IPR011011); BEST Arabidopsis thaliana protein match is: MMD1 (MALE MEIOCYTE DEATH 1); DNA binding / protein binding / zinc ion binding (TAIR:AT1G66170.1); Has 482 Blast hits to 477 proteins in 106 species: Archae - 0; Bacteria - 0; Metazoa - 203; Fungi - 172; Plants - 97; Viruses - 0; Other Eukaryotes - 10 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.3440.1.S1_s_at | 0.03119 | 3.7769 | Glyma13g26960.2 | AT5G12380.1 | annexin, putative; FUNCTIONS IN: calcium-dependent phospholipid binding, calcium ion binding; EXPRESSED IN: embryo, pedicel; EXPRESSED DURING: 4 anthesis, C globular stage; CONTAINS InterPro DOMAIN/s: Annexin like protein (InterPro:IPR015472), Annexin repeat (InterPro:IPR018502), Annexin (InterPro:IPR001464), Annexin repeat, conserved site (InterPro:IPR018252), Annexin, type plant (InterPro:IPR009118); BEST Arabidopsis thaliana protein match is: ANNAT7 (ANNEXIN ARABIDOPSIS 7); calcium ion binding / calcium-dependent phospholipid binding (TAIR:AT5G10230.1); Has 2379 Blast hits to 1795 proteins in 158 species: Archae - 0; Bacteria - 0; Metazoa - 1841; Fungi - 105; Plants - 294; Viruses - 0; Other Eukaryotes - 139 (source: NCBI BLink). |
| GmaAffx.61879.1.S1_at | 0.06742 | 3.5545 | No Soybean Match Identified, E<10E-30, Perecent ID>95% | AT5G60100.1 | Encodes pseudo-response regulator 3 (APRR3/PRR3). PRR3 transcript levels vary in a circadian pattern with peak expression at dusk under long and short day conditions. PRR3 affects the period of the circadian clock and seedlings with reduced levels of PRR3 have shorter periods, based on transcriptional assays of clock-regulated genes. PRR3 is expressed in the vasculature of cotyledons and leaves where it may help stabilize the TOC1 protein by preventing interactions between TOC1 and the F-box protein ZTL. |

| | | | | | |
|---|---|---|---|---|---|
| Gma.11012.1.S1_at | 0.07148 | 3.5111 | Glyma12g31070.1 | AT2G27480.1 | calcium ion binding; FUNCTIONS IN: calcium ion binding; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: stem; CONTAINS InterPro DOMAIN/s: EF-HAND 1 (InterPro:IPR018247), EF-HAND 2 (InterPro:IPR018249), EF-Hand type (InterPro:IPR011992), Calcium-binding EF-hand (InterPro:IPR002048), EF hand (InterPro:IPR018248); BEST Arabidopsis thaliana protein match is: calcium-binding EF hand family protein (TAIR:AT5G04170.1); Has 1279 Blast hits to 1277 proteins in 143 species: Archae - 0; Bacteria - 0; Metazoa - 1048; Fungi - 88; Plants - 73; Viruses - 0; Other Eukaryotes - 70 (source: NCBI BLink). |
| Gma.15381.1.S1_at | 0.04759 | 3.4088 | Glyma18g01340.2 | AT3G48140.1 | senescence-associated protein, putative; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: leaf senescence; LOCATED IN: endomembrane system; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: B12D (InterPro:IPR010530); BEST Arabidopsis thaliana protein match is: germination protein-related (TAIR:AT3G29970.1); Has 78 Blast hits to 78 proteins in 16 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 78; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| Gma.4149.1.S1_at | 0.06962 | 3.361 | Glyma03g31690.1 | AT5G63450.1 | member of CYP94B |
| GmaAffx.92907.1.S1_at | 0.09958 | 3.3432 | Glyma02g01870.1 | AT3G62650.1 | unknown protein; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT2G47485.1); Has 23 Blast hits to 23 proteins in 7 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 23; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.4698.1.S1_at | 0.02653 | 3.1998 | No Soybean Match Identified, E<10E-30, Perecent ID>95% | AT5G60860.1 | Arabidopsis Rab GTPase homolog A1f (AtRABA1f); FUNCTIONS IN: GTP binding; INVOLVED IN: protein transport, small GTPase mediated signal transduction; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Ras GTPase (InterPro:IPR001806), Small GTP-binding protein (InterPro:IPR005225), Ras (InterPro:IPR013753), Ras small GTPase, Rab type (InterPro:IPR003579), Rab11-related (InterPro:IPR015595); BEST Arabidopsis thaliana protein match is: AtRABA1g (Arabidopsis Rab GTPase homolog A1g); GTP binding (TAIR:AT3G15060.1); Has 21819 Blast hits to 21779 proteins in 607 species: Archae - 23; Bacteria - 103; Metazoa - 12072; Fungi - 2821; Plants - 1861; Viruses - 19; Other Eukaryotes - 4920 (source: NCBI BLink). |
| GmaAffx.39545.1.A1_at | 0.03473 | 3.139 | Glyma10g31790.1 | AT1G17820.1 | phosphoinositide binding; FUNCTIONS IN: phosphoinositide binding; INVOLVED IN: signal transduction; LOCATED IN: endomembrane system; CONTAINS InterPro DOMAIN/s: Pleckstrin homology (InterPro:IPR001849); BEST Arabidopsis thaliana protein match is: phosphoinositide binding (TAIR:AT1G73200.1); Has 444 Blast hits to 309 proteins in 108 species: Archae - 0; Bacteria - 7; Metazoa - 132; Fungi - 162; Plants - 27; Viruses - 0; Other Eukaryotes - 116 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.30537.1.S1_at | 0.07712 | 3.1244 | Glyma14g35340.1 | AT4G08950.1 | EXORDIUM (EXO); FUNCTIONS IN: molecular_function unknown; INVOLVED IN: response to brassinosteroid stimulus; LOCATED IN: cell wall, plant-type cell wall; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Phosphate-induced protein 1 conserved region (InterPro:IPR006766); BEST Arabidopsis thaliana protein match is: PHI-1 (PHOSPHATE-INDUCED 1) (TAIR:AT1G35140.1); Has 234 Blast hits to 233 proteins in 17 species: Archae - 0; Bacteria - 2; Metazoa - 0; Fungi - 0; Plants - 232; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| Gma.3081.1.S1_at | 0.02472 | 3.0766 | Glyma15g36180.1 | AT3G12490.2 | Encodes a protein with cysteine proteinase inhibitor activity. Overexpression increases tolerance to abiotic stressors (i.e.salt,osmitic, cold stress). |
| Gma.15530.2.S1_at | 0.0819 | 3.0333 | Glyma01g03540.1 | AT5G05670.1 | signal recognition particle binding; FUNCTIONS IN: signal recognition particle binding; LOCATED IN: endoplasmic reticulum, plasma membrane; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 14 growth stages; BEST Arabidopsis thaliana protein match is: signal recognition particle binding (TAIR:AT2G18770.1); Has 911 Blast hits to 911 proteins in 185 species: Archae - 2; Bacteria - 40; Metazoa - 443; Fungi - 156; Plants - 102; Viruses - 0; Other Eukaryotes - 168 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.93.1.S1_s_at | 0.02679 | 3.0102 | Glyma07g07970.2 | AT2G45820.1 | DNA-binding protein, putative; FUNCTIONS IN: DNA binding; INVOLVED IN: biological_process unknown; LOCATED IN: plasma membrane; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Remorin, C-terminal region (InterPro:IPR005516), Remorin, N-terminal region (InterPro:IPR005518); BEST Arabidopsis thaliana protein match is: DNA-binding family protein / remorin family protein (TAIR:AT3G61260.1); Has 2170 Blast hits to 1469 proteins in 249 species: Archae - 2; Bacteria - 278; Metazoa - 371; Fungi - 179; Plants - 300; Viruses - 4; Other Eukaryotes - 1036 (source: NCBI BLink). |
| Gma.6814.1.S1_at | 0.03947 | 2.8753 | Glyma13g08050.1 | AT5G64310.1 | Encodes arabinogalactan-protein (AGP1). |
| GmaAffx.89235.1.A1_at | 0.06017 | 2.8518 | Glyma01g26840.2 | AT3G54040.1 | photoassimilate-responsive protein-related; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: endomembrane system; EXPRESSED IN: 20 plant structures; EXPRESSED DURING: 10 growth stages; CONTAINS InterPro DOMAIN/s: PAR1 (InterPro:IPR009489); BEST Arabidopsis thaliana protein match is: photoassimilate-responsive protein, putative (TAIR:AT5G52390.1); Has 27 Blast hits to 27 proteins in 6 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 27; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.17929.1.A1_at | 0.07712 | 2.8458 | Glyma04g04260.1 | AT5G42830.1 | transferase family protein; FUNCTIONS IN: transferase activity, transferring acyl groups other than amino-acyl groups, transferase activity; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 17 plant structures; EXPRESSED DURING: 12 growth stages; CONTAINS InterPro DOMAIN/s: Transferase (InterPro:IPR003480); BEST Arabidopsis thaliana protein match is: transferase family protein (TAIR:AT5G07850.1); Has 1313 Blast hits to 1309 proteins in 123 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 86; Plants - 1227; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| Gma.17137.1.S1_at | 0.04307 | 2.8384 | Glyma03g31240.1 | AT2G48130.1 | protease inhibitor/seed storage/lipid transfer protein (LTP) family protein; FUNCTIONS IN: lipid binding; INVOLVED IN: lipid transport; LOCATED IN: anchored to membrane; EXPRESSED IN: hypocotyl, flower, root; EXPRESSED DURING: petal differentiation and expansion stage; CONTAINS InterPro DOMAIN/s: Bifunctional inhibitor/plant lipid transfer protein/seed storage (InterPro:IPR016140), Plant lipid transfer protein/seed storage/trypsin-alpha amylase inhibitor (InterPro:IPR003612), Plant lipid transfer protein/Par allergen (InterPro:IPR000528), Plant lipid transfer protein and hydrophobic protein, helical (InterPro:IPR013770); BEST Arabidopsis thaliana protein match is: protease inhibitor/seed storage/lipid transfer protein (LTP) family protein (TAIR:AT3G22600.1); Has 761 Blast hits to 757 proteins in 56 species: Archae - 0; Bacteria - 6; Metazoa - 1; Fungi - 0; Plants - 752; Viruses - 0; Other Eukaryotes - 2 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.3594.2.S1_a_at | 0.05722 | 2.833 | Glyma17g17970.1 | AT4G37450.1 | AGP18 is a lysine-rich arabinogalactan-protein (AGP) and part of a multi-gene family of glycoproteins with approx. 50 members. It falls into one subclass with AGP17 and AGP19, other lysine-rich AGPs. It is expressed in young leaves, shoots, roots and flowers. |
| GmaAffx.22374.1.S1_at | 0.02679 | 2.7364 | Glyma04g01010.2 | AT2G21300.2 | kinesin motor family protein; FUNCTIONS IN: microtubule motor activity, ATP binding; INVOLVED IN: microtubule-based movement; LOCATED IN: plasma membrane; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 11 growth stages; CONTAINS InterPro DOMAIN/s: Kinesin, motor region (InterPro:IPR001752); BEST Arabidopsis thaliana protein match is: kinesin motor family protein (TAIR:AT4G38950.2); Has 7485 Blast hits to 7163 proteins in 235 species: Archae - 0; Bacteria - 2; Metazoa - 3790; Fungi - 875; Plants - 908; Viruses - 0; Other Eukaryotes - 1910 (source: NCBI BLink). |
| GmaAffx.2557.1.S1_at | 0.04484 | 2.6962 | Glyma05g25660.1 | AT2G34740.1 | catalytic/ protein serine/threonine phosphatase; FUNCTIONS IN: protein serine/threonine phosphatase activity, catalytic activity; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; CONTAINS InterPro DOMAIN/s: Protein phosphatase 2C-related (InterPro:IPR001932), Protein phosphatase 2C (InterPro:IPR015655), Protein phosphatase 2C, N-terminal (InterPro:IPR014045); BEST Arabidopsis thaliana protein match is: protein phosphatase 2C, putative / PP2C, putative (TAIR:AT4G28400.1); Has 5333 Blast hits to 5320 proteins in 675 species: Archae - 7; Bacteria - 1029; Metazoa - 1363; Fungi - 524; Plants - 1327; Viruses - 11; Other Eukaryotes - 1072 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.4331.2.S1_at | 0.06925 | 2.668 | Glyma20g10310.1 | AT3G43720.2 | protease inhibitor/seed storage/lipid transfer protein (LTP) family protein; FUNCTIONS IN: lipid binding; INVOLVED IN: lipid transport; LOCATED IN: anchored to membrane, membrane; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Bifunctional inhibitor/plant lipid transfer protein/seed storage (InterPro:IPR016140), Plant lipid transfer protein/seed storage/trypsin-alpha amylase inhibitor (InterPro:IPR003612); BEST Arabidopsis thaliana protein match is: protease inhibitor/seed storage/lipid transfer protein (LTP) family protein (TAIR:AT2G27130.1); Has 404 Blast hits to 402 proteins in 17 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 404; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| GmaAffx.57046.1.S1_at | 0.04759 | 2.6565 | Glyma11g08540.1 | AT4G35840.1 | zinc finger (C3HC4-type RING finger) family protein; FUNCTIONS IN: protein binding, zinc ion binding; CONTAINS InterPro DOMAIN/s: Zinc finger, RING-type, conserved site (InterPro:IPR017907), Zinc finger, RING-type (InterPro:IPR001841), Zinc finger, RING/FYVE/PHD-type (InterPro:IPR013083); BEST Arabidopsis thaliana protein match is: zinc finger (C3HC4-type RING finger) family protein (TAIR:AT2G17730.1); Has 5941 Blast hits to 5924 proteins in 207 species: Archae - 0; Bacteria - 2; Metazoa - 1867; Fungi - 423; Plants - 2717; Viruses - 21; Other Eukaryotes - 911 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.48718.1.S1_at | 0.05654 | 2.6412 | Glyma12g35410.1 | AT5G53870.1 | plastocyanin-like domain-containing protein; FUNCTIONS IN: electron carrier activity, copper ion binding; LOCATED IN: plasma membrane, anchored to membrane; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 9 growth stages; CONTAINS InterPro DOMAIN/s: Plastocyanin-like (InterPro:IPR003245), Cupredoxin (InterPro:IPR008972); BEST Arabidopsis thaliana protein match is: plastocyanin-like domain-containing protein (TAIR:AT4G27520.1); Has 168884 Blast hits to 72384 proteins in 1909 species: Archae - 291; Bacteria - 24836; Metazoa - 65984; Fungi - 32024; Plants - 14085; Viruses - 4180; Other Eukaryotes - 27484 (source: NCBI BLink). |
| GmaAffx.88221.1.S1_at | 0.07853 | 2.5808 | No Soybean Match Identified, E<10E-30, Perecent ID>95% | AT1G04780.1 | ankyrin repeat family protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: plasma membrane; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Ankyrin (InterPro:IPR002110); BEST Arabidopsis thaliana protein match is: ankyrin repeat family protein (TAIR:AT3G24210.1); Has 832 Blast hits to 578 proteins in 80 species: Archae - 0; Bacteria - 8; Metazoa - 576; Fungi - 12; Plants - 146; Viruses - 2; Other Eukaryotes - 88 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.5267.1.S1_at | 0.09531 | 2.5407 | Glyma15g19120.1 | AT4G25440.1 | zinc finger WD40 repeat protein 1 (ZFWD1); FUNCTIONS IN: zinc ion binding, nucleic acid binding; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: xylem; CONTAINS InterPro DOMAIN/s: Zinc finger, CCCH-type (InterPro:IPR000571), WD40 repeat-like (InterPro:IPR011046), WD40 repeat, region (InterPro:IPR017986), WD40/YVTN repeat-like (InterPro:IPR015943), WD40 repeat (InterPro:IPR001680); BEST Arabidopsis thaliana protein match is: WD-40 repeat family protein / zfwd2 protein (ZFWD2), putative (TAIR:AT5G51980.1); Has 20982 Blast hits to 11831 proteins in 417 species: Archae - 40; Bacteria - 3055; Metazoa - 9003; Fungi - 4181; Plants - 1744; Viruses - 0; Other Eukaryotes - 2959 (source: NCBI BLink). |
| Gma.4531.2.S1_at | 0.05233 | 2.508 | Glyma06g10710.1 | AT4G08950.1 | EXORDIUM (EXO); FUNCTIONS IN: molecular_function unknown; INVOLVED IN: response to brassinosteroid stimulus; LOCATED IN: cell wall, plant-type cell wall; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Phosphate-induced protein 1 conserved region (InterPro:IPR006766); BEST Arabidopsis thaliana protein match is: PHI-1 (PHOSPHATE-INDUCED 1) (TAIR:AT1G35140.1); Has 234 Blast hits to 233 proteins in 17 species: Archae - 0; Bacteria - 2; Metazoa - 0; Fungi - 0; Plants - 232; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.3440.2.S1_at | 0.0503 | 2.4909 | Glyma15g38010.2 | AT5G12380.1 | annexin, putative; FUNCTIONS IN: calcium-dependent phospholipid binding, calcium ion binding; EXPRESSED IN: embryo, pedicel; EXPRESSED DURING: 4 anthesis, C globular stage; CONTAINS InterPro DOMAIN/s: Annexin like protein (InterPro:IPR015472), Annexin repeat (InterPro:IPR018502), Annexin (InterPro:IPR001464), Annexin repeat, conserved site (InterPro:IPR018252), Annexin, type plant (InterPro:IPR009118); BEST Arabidopsis thaliana protein match is: ANNAT7 (ANNEXIN ARABIDOPSIS 7); calcium ion binding / calcium-dependent phospholipid binding (TAIR:AT5G10230.1); Has 2379 Blast hits to 1795 proteins in 158 species: Archae - 0; Bacteria - 0; Metazoa - 1841; Fungi - 105; Plants - 294; Viruses - 0; Other Eukaryotes - 139 (source: NCBI BLink). |
| GmaAffx.52628.1.S1_at | 0.09793 | 2.4271 | Glyma17g03960.1 | AT4G16280.2 | Involved in the promotion of the transition of the vegetative meristem to reproductive development. Four forms of the protein (alpha, beta, delta and gamma) are produced by alternative splicing. Involved in RNA-mediated chromatin silencing. |
| Gma.13844.1.A1_at | 0.09858 | 2.4002 | Glyma08g39250.1 | AT5G05670.1 | signal recognition particle binding; FUNCTIONS IN: signal recognition particle binding; LOCATED IN: endoplasmic reticulum, plasma membrane; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 14 growth stages; BEST Arabidopsis thaliana protein match is: signal recognition particle binding (TAIR:AT2G18770.1); Has 911 Blast hits to 911 proteins in 185 species: Archae - 2; Bacteria - 40; Metazoa - 443; Fungi - 156; Plants - 102; Viruses - 0; Other Eukaryotes - 168 (source: NCBI BLink). |
| GmaAffx.19751.1.A1_at | 0.07755 | 2.3747 | Glyma02g41140.1 | AT5G37660.1 | Encodes a plasmodesmal protein that may be involved in the intercellular movement of molecules through the plasmodesmata. The protein has two DUF26 domains and a single transmembrane domain. |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.51587.1.S1_at | 0.05948 | 2.3667 | No Soybean Match Identified, E<10E-30, Perecent ID>95% | AT5G54650.2 | Encodes a protein with similarity to formins that is involved in cytokinesis. Loss of function mutations exhibit delayed cellularization during endosperm development. FH5 is expressed in the endosperm and the protein localizes to the cell plate. FH5 was shown to be a maternally expressed imprinted gene. |
| GmaAffx.45036.1.S1_at | 0.04466 | 2.3551 | Glyma19g04410.1 | AT5G15780.1 | pollen Ole e 1 allergen and extensin family protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: endomembrane system; EXPRESSED IN: 14 plant structures; EXPRESSED DURING: 10 growth stages; CONTAINS InterPro DOMAIN/s: Pollen Ole e 1 allergen and extensin (InterPro:IPR006041); Has 75182 Blast hits to 31819 proteins in 1369 species: Archae - 252; Bacteria - 13906; Metazoa - 31718; Fungi - 6983; Plants - 10029; Viruses - 2694; Other Eukaryotes - 9600 (source: NCBI BLink). |
| Gma.10717.2.S1_s_at | 0.03119 | 2.339 | Glyma17g02600.2 | AT3G06390.1 | integral membrane family protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: root; CONTAINS InterPro DOMAIN/s: Uncharacterised protein family UPF0497, trans-membrane plant subgroup (InterPro:IPR006702), Uncharacterised protein family UPF0497, trans-membrane plant subgroup (InterPro:IPR006459); BEST Arabidopsis thaliana protein match is: integral membrane family protein (TAIR:AT4G15610.1); Has 278 Blast hits to 278 proteins in 13 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 278; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.14100.1.S1_at | 0.09403 | 2.3158 | Glyma11g37780.1 | AT1G67856.1 | protein binding / zinc ion binding; FUNCTIONS IN: protein binding, zinc ion binding; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; CONTAINS InterPro DOMAIN/s: Zinc finger, RING-type (InterPro:IPR001841), Zinc finger, RING/FYVE/PHD-type (InterPro:IPR013083); BEST Arabidopsis thaliana protein match is: zinc finger (C3HC4-type RING finger) family protein (TAIR:AT1G24580.1); Has 3925 Blast hits to 3916 proteins in 187 species: Archae - 0; Bacteria - 0; Metazoa - 1142; Fungi - 305; Plants - 1895; Viruses - 19; Other Eukaryotes - 564 (source: NCBI BLink). |
| Gma.6850.1.S1_at | 0.06293 | 2.2838 | Glyma04g04240.1 | AT5G42830.1 | transferase family protein; FUNCTIONS IN: transferase activity, transferring acyl groups other than amino-acyl groups, transferase activity; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 17 plant structures; EXPRESSED DURING: 12 growth stages; CONTAINS InterPro DOMAIN/s: Transferase (InterPro:IPR003480); BEST Arabidopsis thaliana protein match is: transferase family protein (TAIR:AT5G07850.1); Has 1313 Blast hits to 1309 proteins in 123 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 86; Plants - 1227; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.36779.1.S1_at | 0.08725 | 2.2232 | Glyma20g30860.1 | AT4G24630.1 | receptor/ zinc ion binding; FUNCTIONS IN: receptor activity, zinc ion binding; INVOLVED IN: multicellular organismal development; LOCATED IN: plasma membrane; EXPRESSED IN: 12 plant structures; EXPRESSED DURING: 8 growth stages; CONTAINS InterPro DOMAIN/s: Zinc finger, DHHC-type (InterPro:IPR001594), Plexin-like fold (InterPro:IPR016201); BEST Arabidopsis thaliana protein match is: zinc finger (DHHC type) family protein (TAIR:AT5G50020.1); Has 3990 Blast hits to 3983 proteins in 185 species: Archae - 0; Bacteria - 0; Metazoa - 1956; Fungi - 539; Plants - 407; Viruses - 0; Other Eukaryotes - 1088 (source: NCBI BLink). |
| GmaAffx.90113.1.S1_at | 0.02679 | 2.198 | Ambiguous Hit | AT1G52690.2 | late embryogenesis abundant protein, putative / LEA protein, putative; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: embryonic development ending in seed dormancy; LOCATED IN: cellular_component unknown; EXPRESSED IN: 9 plant structures; EXPRESSED DURING: dry seed stage, 4 anthesis, LP.04 four leaves visible, petal differentiation and expansion stage; BEST Arabidopsis thaliana protein match is: late embryogenesis abundant protein, putative / LEA protein, putative (TAIR:AT3G15670.1); Has 2025 Blast hits to 947 proteins in 265 species: Archae - 6; Bacteria - 474; Metazoa - 312; Fungi - 113; Plants - 812; Viruses - 15; Other Eukaryotes - 293 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.59136.1.A1_at | 0.08428 | 2.1828 | Glyma02g01860.1 | AT2G30620.1 | histone H1.2; FUNCTIONS IN: DNA binding; INVOLVED IN: nucleosome assembly; LOCATED IN: nucleolus, nucleus, nucleosome; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Winged helix repressor DNA-binding (InterPro:IPR011991), Histone H5 (InterPro:IPR005819), Histone H1/H5 (InterPro:IPR005818); BEST Arabidopsis thaliana protein match is: histone H1, putative (TAIR:AT1G06760.1); Has 17371 Blast hits to 9398 proteins in 1065 species: Archae - 39; Bacteria - 6415; Metazoa - 4684; Fungi - 1032; Plants - 746; Viruses - 137; Other Eukaryotes - 4318 (source: NCBI BLink). |
| Gma.8704.1.S1_at | 0.09985 | 2.1459 | Glyma06g10970.1 | AT1G78120.1 | tetratricopeptide repeat (TPR)-containing protein; FUNCTIONS IN: binding; LOCATED IN: mitochondrion; EXPRESSED IN: 9 plant structures; EXPRESSED DURING: 4 anthesis, petal differentiation and expansion stage; CONTAINS InterPro DOMAIN/s: Molecular chaperone, heat shock protein, Hsp40, DnaJ (InterPro:IPR015609), Tetratricopeptide TPR-1 (InterPro:IPR001440), Tetratricopeptide-like helical (InterPro:IPR011990), Tetratricopeptide TPR2 (InterPro:IPR013105), Tetratricopeptide region (InterPro:IPR013026); BEST Arabidopsis thaliana protein match is: tetratricopeptide repeat (TPR)-containing protein (TAIR:AT5G65160.1); Has 11653 Blast hits to 7027 proteins in 589 species: Archae - 508; Bacteria - 2639; Metazoa - 3642; Fungi - 1070; Plants - 906; Viruses - 0; Other Eukaryotes - 2888 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.13245.1.S1_at | 0.09647 | 2.1448 | Glyma15g31290.1 | AT4G25150.1 | acid phosphatase, putative; FUNCTIONS IN: acid phosphatase activity; INVOLVED IN: biological_process unknown; LOCATED IN: endomembrane system; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 9 growth stages; CONTAINS InterPro DOMAIN/s: Acid phosphatase (Class B) (InterPro:IPR005519), Vegetative storage protein/acid phosphatase (InterPro:IPR014403), Acid phosphatase, plant (InterPro:IPR010028); BEST Arabidopsis thaliana protein match is: acid phosphatase, putative (TAIR:AT5G51260.1); Has 559 Blast hits to 559 proteins in 162 species: Archae - 0; Bacteria - 269; Metazoa - 2; Fungi - 0; Plants - 241; Viruses - 0; Other Eukaryotes - 47 (source: NCBI BLink). |
| Gma.15570.1.S1_s_at | 0.07224 | 2.1342 | Glyma18g17580.1 | AT1G13750.1 | calcineurin-like phosphoesterase family protein; FUNCTIONS IN: hydrolase activity, protein serine/threonine phosphatase activity, metal ion binding, acid phosphatase activity; INVOLVED IN: biological_process unknown; LOCATED IN: endomembrane system; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Purple acid phosphatase, N-terminal (InterPro:IPR015914), Metallophosphoesterase (InterPro:IPR004843), Fibronectin, type III (InterPro:IPR003961), Purple acid phosphatase-like, N-terminal (InterPro:IPR008963); BEST Arabidopsis thaliana protein match is: PAP27 (PURPLE ACID PHOSPHATASE 27); acid phosphatase/ protein serine/threonine phosphatase (TAIR:AT5G50400.1); Has 1113 Blast hits to 1105 proteins in 225 species: Archae - 0; Bacteria - 236; Metazoa - 176; Fungi - 60; Plants - 415; Viruses - 0; Other Eukaryotes - 226 (source: NCBI BLink). |
| Gma.7140.1.S1_at | 0.0687 | 2.1158 | Glyma10g34220.2 | AT5G04430.2 | Gene model AT5G04430.1 produces active protein. (BTS1S). Binds to ToMV genomic RNA and prevents viral multiplication |

| | | | | | |
|---|---|---|---|---|---|
| Gma.2590.7.S1_at | 0.06637 | 2.0696 | Ambiguous Hit | AT2G23810.1 | Member of TETRASPANIN family |
| Gma.6274.1.S1_at | 0.06847 | 2.0309 | Glyma04g05270.1 | AT5G11740.1 | Encodes arabinogalactan protein (AGP15). |
| GmaAffx.67038.1.S1_at | 0.06438 | 2.0247 | Glyma06g03960.1 | AT1G47710.1 | serpin, putative / serine protease inhibitor, putative; FUNCTIONS IN: serine-type endopeptidase inhibitor activity, cysteine-type endopeptidase inhibitor activity; INVOLVED IN: biological_process unknown; LOCATED IN: apoplast; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Protease inhibitor I4, serpin, plant (InterPro:IPR015554), Protease inhibitor I4, serpin (InterPro:IPR000215); BEST Arabidopsis thaliana protein match is: serpin, putative / serine protease inhibitor, putative (TAIR:AT3G45220.1); Has 4823 Blast hits to 4760 proteins in 347 species: Archae - 52; Bacteria - 220; Metazoa - 3736; Fungi - 1; Plants - 230; Viruses - 419; Other Eukaryotes - 165 (source: NCBI BLink). |
| GmaAffx.11131.1.A1_at | 0.04307 | 2.0032 | Glyma19g38460.1 | AT5G22250.1 | CCR4-NOT transcription complex protein, putative; FUNCTIONS IN: ribonuclease activity, nucleic acid binding; INVOLVED IN: RNA modification; LOCATED IN: nucleus; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 11 growth stages; CONTAINS InterPro DOMAIN/s: Ribonuclease CAF1 (InterPro:IPR006941), Polynucleotidyl transferase, Ribonuclease H fold (InterPro:IPR012337); BEST Arabidopsis thaliana protein match is: CCR4-NOT transcription complex protein, putative (TAIR:AT3G44260.1); Has 630 Blast hits to 622 proteins in 157 species: Archae - 0; Bacteria - 0; Metazoa - 218; Fungi - 95; Plants - 219; Viruses - 0; Other Eukaryotes - 98 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.4240.1.S1_at | 0.0687 | 2.0014 | Glyma13g19440.1 | AT2G37520.1 | PHD finger family protein; FUNCTIONS IN: protein binding, DNA binding, zinc ion binding; INVOLVED IN: regulation of transcription, DNA-dependent; LOCATED IN: nucleus; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Zinc finger, RING/FYVE/PHD-type (InterPro:IPR013083), Zinc finger, PHD-type (InterPro:IPR001965), Zinc finger, FYVE/PHD-type (InterPro:IPR011011), Acyl-CoA N-acyltransferase (InterPro:IPR016181); BEST Arabidopsis thaliana protein match is: PHD finger transcription factor, putative (TAIR:AT3G53680.1); Has 2686 Blast hits to 2189 proteins in 145 species: Archae - 0; Bacteria - 0; Metazoa - 1957; Fungi - 205; Plants - 372; Viruses - 0; Other Eukaryotes - 152 (source: NCBI BLink). |
| GmaAffx.33712.2.S1_at | 0.08261 | 1.9891 | Glyma01g39810.1 | AT4G37450.1 | AGP18 is a lysine-rich arabinogalactan-protein (AGP) and part of a multi-gene family of glycoproteins with approx. 50 members. It falls into one subclass with AGP17 and AGP19, other lysine-rich AGPs. It is expressed in young leaves, shoots, roots and flowers. |
| Gma.8149.1.S1_at | 0.06347 | 1.9829 | Glyma14g10590.1 | AT5G11740.1 | Encodes arabinogalactan protein (AGP15). |
| Gma.3594.1.S1_at | 0.03369 | 1.9752 | Glyma05g21820.1 | AT4G37450.1 | AGP18 is a lysine-rich arabinogalactan-protein (AGP) and part of a multi-gene family of glycoproteins with approx. 50 members. It falls into one subclass with AGP17 and AGP19, other lysine-rich AGPs. It is expressed in young leaves, shoots, roots and flowers. |

| | | | | | |
|---|---|---|---|---|---|
| Gma.10888.2.A1_at | 0.06238 | 1.9638 | Glyma17g29800.2 | AT5G25560.1 | zinc finger (C3HC4-type RING finger) family protein; FUNCTIONS IN: protein binding, zinc ion binding; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Zinc finger, CTCHY-type (InterPro:IPR017921), Zinc finger, CHY-type (InterPro:IPR008913), Zinc finger, RING-type (InterPro:IPR001841), Zinc finger, RING/FYVE/PHD-type (InterPro:IPR013083); BEST Arabidopsis thaliana protein match is: zinc finger (C3HC4-type RING finger) family protein (TAIR:AT5G22920.1); Has 631 Blast hits to 629 proteins in 118 species: Archae - 2; Bacteria - 0; Metazoa - 274; Fungi - 89; Plants - 162; Viruses - 0; Other Eukaryotes - 104 (source: NCBI BLink). |
| GmaAffx.72630.1.S1_at | 0.0951 | 1.9609 | Glyma14g34640.1 | AT5G65160.1 | tetratricopeptide repeat (TPR)-containing protein; FUNCTIONS IN: binding; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 6 plant structures; EXPRESSED DURING: L mature pollen stage, M germinated pollen stage, 4 anthesis, petal differentiation and expansion stage; CONTAINS InterPro DOMAIN/s: Molecular chaperone, heat shock protein, Hsp40, DnaJ (InterPro:IPR015609), Tetratricopeptide TPR-1 (InterPro:IPR001440), Tetratricopeptide-like helical (InterPro:IPR011990), Tetratricopeptide TPR2 (InterPro:IPR013105), Tetratricopeptide region (InterPro:IPR013026); BEST Arabidopsis thaliana protein match is: tetratricopeptide repeat (TPR)-containing protein (TAIR:AT5G10090.1); Has 12195 Blast hits to 7361 proteins in 598 species: Archae - 415; Bacteria - 2850; Metazoa - 3545; Fungi - 998; Plants - 998; Viruses - 0; Other Eukaryotes - 3389 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.8337.1.S1_a_at | 0.09985 | 1.9463 | No Soybean Match Identified, E<10E-30, Perecent ID>95% | AT2G18090.1 | PHD finger family protein / SWIB complex BAF60b domain-containing protein / GYF domain-containing protein; FUNCTIONS IN: protein binding, DNA binding, zinc ion binding; INVOLVED IN: regulation of transcription, DNA-dependent; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: SWIB/MDM2 (InterPro:IPR003121), Zinc finger, PHD-type (InterPro:IPR001965), Zinc finger, FYVE/PHD-type (InterPro:IPR011011), GYF (InterPro:IPR003169); BEST Arabidopsis thaliana protein match is: DNA binding / nucleic acid binding / protein binding / zinc ion binding (TAIR:AT3G51120.1); Has 879 Blast hits to 782 proteins in 136 species: Archae - 0; Bacteria - 105; Metazoa - 439; Fungi - 39; Plants - 214; Viruses - 0; Other Eukaryotes - 82 (source: NCBI BLink). |
| Gma.17642.1.S1_at | 0.0505 | 1.9401 | Glyma18g10070.1 | AT1G04970.1 | lipid-binding serum glycoprotein family protein; FUNCTIONS IN: lipid binding; INVOLVED IN: biological_process unknown; LOCATED IN: plasma membrane, vacuole; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Bactericidal permeability-increasing protein, alpha/beta domain (InterPro:IPR017943), Lipid-binding serum glycoprotein, N-terminal (InterPro:IPR017942), Lipid-binding serum glycoprotein, C-terminal (InterPro:IPR001124); BEST Arabidopsis thaliana protein match is: lipid-binding serum glycoprotein family protein (TAIR:AT3G20270.2); Has 354 Blast hits to 349 proteins in 47 species: Archae - 0; Bacteria - 0; Metazoa - 309; Fungi - 0; Plants - 26; Viruses - 0; Other Eukaryotes - 19 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.16921.1.S1_at | 0.0881 | 1.938 | Glyma20g26770.1 | AT4G33920.1 | protein phosphatase 2C family protein / PP2C family protein; FUNCTIONS IN: protein serine/threonine phosphatase activity, catalytic activity; INVOLVED IN: protein amino acid dephosphorylation; LOCATED IN: mitochondrion, protein serine/threonine phosphatase complex; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Protein phosphatase 2C, manganese/magnesium aspartate binding site (InterPro:IPR000222), Protein phosphatase 2C-related (InterPro:IPR001932), Protein phosphatase 2C (InterPro:IPR015655), Protein phosphatase 2C, N-terminal (InterPro:IPR014045); BEST Arabidopsis thaliana protein match is: protein phosphatase 2C, putative / PP2C, putative (TAIR:AT3G51370.1); Has 3521 Blast hits to 3519 proteins in 216 species: Archae - 0; Bacteria - 7; Metazoa - 1157; Fungi - 386; Plants - 1254; Viruses - 3; Other Eukaryotes - 714 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.1494.1.S1_x_at | 0.09907 | 1.9314 | Glyma17g11940.1 | AT2G10940.2 | protease inhibitor/seed storage/lipid transfer protein (LTP) family protein; FUNCTIONS IN: lipid binding; INVOLVED IN: lipid transport; LOCATED IN: chloroplast thylakoid membrane, apoplast, chloroplast, membrane; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Bifunctional inhibitor/plant lipid transfer protein/seed storage (InterPro:IPR016140), Plant lipid transfer protein/seed storage/trypsin-alpha amylase inhibitor (InterPro:IPR003612), Plant lipid transfer protein and hydrophobic protein, helical (InterPro:IPR013770); BEST Arabidopsis thaliana protein match is: protease inhibitor/seed storage/lipid transfer protein (LTP) family protein (TAIR:AT1G62500.1); Has 41695 Blast hits to 15869 proteins in 1031 species: Archae - 182; Bacteria - 8355; Metazoa - 15253; Fungi - 3374; Plants - 6727; Viruses - 1536; Other Eukaryotes - 6268 (source: NCBI BLink). |
| Gma.2331.3.A1_at | 0.04783 | 1.9173 | Glyma06g09810.1 | AT5G49700.1 | DNA-binding protein-related; INVOLVED IN: biological_process unknown; EXPRESSED IN: shoot apex, embryo, inflorescence meristem; EXPRESSED DURING: D bilateral stage; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF296 (InterPro:IPR005175), Predicted AT-hook DNA-binding (InterPro:IPR014476); BEST Arabidopsis thaliana protein match is: DNA-binding protein-related (TAIR:AT1G14490.1); Has 424 Blast hits to 423 proteins in 20 species: Archae - 0; Bacteria - 2; Metazoa - 3; Fungi - 2; Plants - 417; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.4549.2.S1_at | 0.05811 | 1.9085 | Glyma04g06580.2 | AT2G25110.1 | Encodes an endoplasmic reticulum protein SDF2 (stromal-derived factor-2). Forms a complex SDF2-ERdj3B-BiP that is required for the proper accumulation of the surface-exposed leucine-rich repeat receptor kinases EFR. EFR is involved in PAMP (pathogen associated molecular patterns) triggered immunity. |
| Gma.6411.1.S1_at | 0.05196 | 1.908 | Glyma02g18320.1 | AT1G71950.1 | identical protein binding / serine-type endopeptidase; FUNCTIONS IN: identical protein binding, serine-type endopeptidase activity; INVOLVED IN: negative regulation of catalytic activity; LOCATED IN: plasma membrane; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Proteinase inhibitor, propeptide (InterPro:IPR009020), Proteinase inhibitor I9, subtilisin propeptide (InterPro:IPR010259), Peptidase S8, subtilisin-related (InterPro:IPR015500); BEST Arabidopsis thaliana protein match is: subtilase family protein (TAIR:AT1G32950.1); Has 441 Blast hits to 438 proteins in 21 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 441; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.6031.1.A1_at | 0.07119 | 1.893 | No Soybean Match Identified, E<10E-30, Perecent ID>95% | AT3G12640.1 | RNA binding / nucleic acid binding / nucleotide binding; FUNCTIONS IN: RNA binding, nucleotide binding, nucleic acid binding; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: RNA recognition motif, RNP-1 (InterPro:IPR000504), Nucleotide-binding, alpha-beta plait (InterPro:IPR012677); BEST Arabidopsis thaliana protein match is: RNA recognition motif (RRM)-containing protein (TAIR:AT2G24350.1); Has 5419 Blast hits to 5243 proteins in 418 species: Archae - 0; Bacteria - 543; Metazoa - 2418; Fungi - 873; Plants - 1071; Viruses - 1; Other Eukaryotes - 513 (source: NCBI BLink). |
| GmaAffx.82962.1.S1_at | 0.09268 | 1.8868 | Glyma03g37140.1 | AT2G47770.1 | Encodes a membrane-bound protein designated AtTSPO (Arabidopsis thaliana TSPO-related). AtTSPO is related to the bacterial outer membrane tryptophan-rich sensory protein (TspO) and the mammalian mitochondrial 18 kDa Translocator Protein (18 kDa TSPO), members of the TspO/MBR domain-containing membrane proteins. Mainly detected in dry seeds, but can be induced in vegetative tissues by osmotic or salt stress or abscisic acid treatment. Located in endoplasmic reticulum and the Golgi stacks. |

| | | | | | |
|---|---|---|---|---|---|
| Gma.6596.1.S1_at | 0.03237 | 1.8574 | Glyma09g26610.2 | AT5G07250.1 | ARABIDOPSIS RHOMBOID-LIKE PROTEIN 3 (ATRBL3); FUNCTIONS IN: molecular_function unknown; LOCATED IN: integral to membrane; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Peptidase S54, rhomboid (InterPro:IPR002610); BEST Arabidopsis thaliana protein match is: ATRBL2 (Arabidopsis thaliana Rhomboid-like 2); serine-type endopeptidase (TAIR:AT1G63120.1); Has 4072 Blast hits to 4072 proteins in 965 species: Archae - 91; Bacteria - 2083; Metazoa - 468; Fungi - 106; Plants - 207; Viruses - 0; Other Eukaryotes - 1117 (source: NCBI BLink). |
| Gma.89.1.S1_at | 0.06018 | 1.8517 | Glyma06g15360.2 | AT1G43130.1 | LIKE COV 2 (LCV2); FUNCTIONS IN: molecular_function unknown; INVOLVED IN: stem vascular tissue pattern formation; LOCATED IN: cellular_component unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF502 (InterPro:IPR007462); BEST Arabidopsis thaliana protein match is: COV1 (CONTINUOUS VASCULAR RING) (TAIR:AT2G20120.1); Has 1946 Blast hits to 1946 proteins in 369 species: Archae - 4; Bacteria - 692; Metazoa - 0; Fungi - 0; Plants - 84; Viruses - 0; Other Eukaryotes - 1166 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.1337.1.S1_s_at | 0.06299 | 1.8504 | Glyma05g02630.1 | AT4G24690.1 | ubiquitin-associated (UBA)/TS-N domain-containing protein / octicosapeptide/Phox/Bem1 (PB1) domain-containing protein; FUNCTIONS IN: zinc ion binding; INVOLVED IN: biological_process unknown; LOCATED IN: cytoplasm; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Octicosapeptide/Phox/Bem1p (InterPro:IPR000270), Ubiquitin-associated/translation elongation factor EF1B, N-terminal (InterPro:IPR000449), Zinc finger, ZZ-type (InterPro:IPR000433); Has 492 Blast hits to 472 proteins in 101 species: Archae - 0; Bacteria - 6; Metazoa - 329; Fungi - 68; Plants - 57; Viruses - 0; Other Eukaryotes - 32 (source: NCBI BLink). |
| Gma.15699.1.S1_at | 0.08232 | 1.829 | Glyma08g14570.1 | AT4G39740.1 | electron transport SCO1/SenC family protein; FUNCTIONS IN: molecular_function unknown; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Copper chaperone SCO1/SenC (InterPro:IPR003782), Thioredoxin fold (InterPro:IPR012335), Thioredoxin-like fold (InterPro:IPR012336); BEST Arabidopsis thaliana protein match is: electron transport SCO1/SenC family protein (TAIR:AT3G08950.1); Has 2451 Blast hits to 2451 proteins in 612 species: Archae - 9; Bacteria - 1265; Metazoa - 132; Fungi - 105; Plants - 35; Viruses - 0; Other Eukaryotes - 905 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.5804.1.S1_at | 0.05233 | 1.8279 | Ambiguous Hit | AT5G58160.1 | actin binding; FUNCTIONS IN: actin binding; INVOLVED IN: cellular component organization, actin cytoskeleton organization; EXPRESSED IN: 8 plant structures; EXPRESSED DURING: 4 anthesis, petal differentiation and expansion stage; CONTAINS InterPro DOMAIN/s: Actin-binding FH2 and DRF autoregulatory (InterPro:IPR003104), C2 calcium/lipid-binding region, CaLB (InterPro:IPR008973), Actin-binding FH2 (InterPro:IPR015425), Tensin phosphatase, C2 domain (InterPro:IPR014020); BEST Arabidopsis thaliana protein match is: formin homology 2 domain-containing protein / FH2 domain-containing protein (TAIR:AT2G25050.1); Has 176085 Blast hits to 80584 proteins in 2071 species: Archae - 318; Bacteria - 30726; Metazoa - 76316; Fungi - 24501; Plants - 18155; Viruses - 5055; Other Eukaryotes - 21014 (source: NCBI BLink). |
| GmaAffx.29126.1.S1_at | 0.07717 | 1.816 | Glyma03g34080.1 | AT2G36910.1 | Belongs to the family of ATP-binding cassette (ABC) transporters. Also known as AtMDR1.Possibly regulates auxin-dependent responses by influencing basipetal auxin transport in the root. Exerts nonredundant, partially overlapping functions with the ABC transporter encoded by AT3G28860. PGP1 mediates cellular efflux of IAA and interacts with PIN genes that may confer an accelerated vectoral component to PGP-mediated transport. The non-polar localization of PGP1 at root and shoot apices, where IAA gradient-driven transport is impaired, may be required to confer directionality to auxin transport in those tissues. |

| | | | | | |
|---|---|---|---|---|---|
| Gma.17889.1.S1_at | 0.0963 | 1.8127 | Glyma17g09900.1 | AT1G24450.1 | NUCLEAR FUSION DEFECTIVE 2 (NFD2); FUNCTIONS IN: RNA binding, ribonuclease III activity; INVOLVED IN: RNA processing; LOCATED IN: vacuole; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Ribonuclease III (InterPro:IPR000999); Has 1405 Blast hits to 1405 proteins in 451 species: Archae - 3; Bacteria - 885; Metazoa - 2; Fungi - 0; Plants - 7; Viruses - 0; Other Eukaryotes - 508 (source: NCBI BLink). |
| GmaAffx.80393.1.A1_at | 0.0448 | 1.8095 | Glyma13g06860.1 | AT5G15780.1 | pollen Ole e 1 allergen and extensin family protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: endomembrane system; EXPRESSED IN: 14 plant structures; EXPRESSED DURING: 10 growth stages; CONTAINS InterPro DOMAIN/s: Pollen Ole e 1 allergen and extensin (InterPro:IPR006041); Has 75182 Blast hits to 31819 proteins in 1369 species: Archae - 252; Bacteria - 13906; Metazoa - 31718; Fungi - 6983; Plants - 10029; Viruses - 2694; Other Eukaryotes - 9600 (source: NCBI BLink). |
| GmaAffx.34626.1.S1_at | 0.05603 | 1.8058 | Glyma14g36800.1 | AT4G01575.1 | serine protease inhibitor, Kazal-type family protein; FUNCTIONS IN: serine-type endopeptidase inhibitor activity; INVOLVED IN: biological_process unknown; LOCATED IN: endomembrane system; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Proteinase inhibitor I1, Kazal (InterPro:IPR002350); BEST Arabidopsis thaliana protein match is: serine protease inhibitor, Kazal-type family protein (TAIR:AT3G61980.1); Has 23 Blast hits to 23 proteins in 7 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 23; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.8273.1.A1_at | 0.06494 | 1.8001 | Glyma17g33930.1 | AT2G25050.1 | formin homology 2 domain-containing protein / FH2 domain-containing protein; FUNCTIONS IN: actin binding; INVOLVED IN: cellular component organization, actin cytoskeleton organization; CONTAINS InterPro DOMAIN/s: Phosphatase tensin type (InterPro:IPR014019), Actin-binding FH2 and DRF autoregulatory (InterPro:IPR003104), C2 calcium/lipid-binding region, CaLB (InterPro:IPR008973), Actin-binding FH2 (InterPro:IPR015425), Tensin phosphatase, C2 domain (InterPro:IPR014020); BEST Arabidopsis thaliana protein match is: formin homology 2 domain-containing protein / FH2 domain-containing protein (TAIR:AT3G32400.1); Has 81727 Blast hits to 43952 proteins in 1298 species: Archae - 142; Bacteria - 9305; Metazoa - 36688; Fungi - 12327; Plants - 9705; Viruses - 2404; Other Eukaryotes - 11156 (source: NCBI BLink). |
| GmaAffx.43631.1.S1_at | 0.04164 | 1.7936 | Glyma17g09210.1 | AT4G24690.1 | |
| Gma.400.2.S1_at | 0.06291 | 1.7895 | Glyma03g17870.2 | AT3G09350.1 | armadillo/beta-catenin repeat family protein; FUNCTIONS IN: binding; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Armadillo-like helical (InterPro:IPR011989), Armadillo-type fold (InterPro:IPR016024); BEST Arabidopsis thaliana protein match is: armadillo/beta-catenin repeat family protein (TAIR:AT3G53800.1); Has 443 Blast hits to 439 proteins in 147 species: Archae - 0; Bacteria - 178; Fungi - 109; Plants - 94; Viruses - 0; Other Eukaryotes - 62 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.25713.1.S1_at | 0.05778 | 1.7719 | Glyma13g20730.1 | AT5G03330.2 | OTU-like cysteine protease family protein; FUNCTIONS IN: cysteine-type peptidase activity; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; CONTAINS InterPro DOMAIN/s: Ovarian tumour, otubain (InterPro:IPR003323); BEST Arabidopsis thaliana protein match is: OTU-like cysteine protease family protein (TAIR:AT5G04250.2); Has 747 Blast hits to 732 proteins in 134 species: Archae - 0; Bacteria - 4; Metazoa - 363; Fungi - 65; Plants - 143; Viruses - 13; Other Eukaryotes - 159 (source: NCBI BLink). |
| Gma.12765.1.A1_at | 0.07082 | 1.7388 | Glyma04g06430.1 | AT4G31820.1 | A member of the NPY family genes (NPY1/AT4G31820, NPY2/AT2G14820, NPY3/AT5G67440, NPY4/AT2G23050, NPY5/AT4G37590). Encodes a protein with similarity to NHP3. Contains BTB/POZ domain. Promoter region has canonical auxin response element binding site and Wus binding site. Co-localizes to the late endosome with PID. Regulates cotyledon development through control of PIN1 polarity in concert with PID. Also involved in sepal and gynoecia development. |

| | | | | | |
|---|---|---|---|---|---|
| Gma.3661.1.S1_s_at | 0.09797 | 1.7205 | Glyma04g03860.1 | AT1G47710.1 | serpin, putative / serine protease inhibitor, putative; FUNCTIONS IN: serine-type endopeptidase inhibitor activity, cysteine-type endopeptidase inhibitor activity; INVOLVED IN: biological_process unknown; LOCATED IN: apoplast; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Protease inhibitor I4, serpin, plant (InterPro:IPR015554), Protease inhibitor I4, serpin (InterPro:IPR000215); BEST Arabidopsis thaliana protein match is: serpin, putative / serine protease inhibitor, putative (TAIR:AT3G45220.1); Has 4823 Blast hits to 4760 proteins in 347 species: Archae - 52; Bacteria - 220; Metazoa - 3736; Fungi - 1; Plants - 230; Viruses - 419; Other Eukaryotes - 165 (source: NCBI BLink). |
| Gma.17756.1.A1_s_at | 0.05678 | 1.7139 | Glyma08g24840.1 | AT4G25150.1 | acid phosphatase, putative; FUNCTIONS IN: acid phosphatase activity; INVOLVED IN: biological_process unknown; LOCATED IN: endomembrane system; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 9 growth stages; CONTAINS InterPro DOMAIN/s: Acid phosphatase (Class B) (InterPro:IPR005519), Vegetative storage protein/acid phosphatase (InterPro:IPR014403), Acid phosphatase, plant (InterPro:IPR010028); BEST Arabidopsis thaliana protein match is: acid phosphatase, putative (TAIR:AT5G51260.1); Has 559 Blast hits to 559 proteins in 162 species: Archae - 0; Bacteria - 269; Metazoa - 2; Fungi - 0; Plants - 241; Viruses - 0; Other Eukaryotes - 47 (source: NCBI BLink). |
| GmaAffx.16446.1.S1_at | 0.06877 | 1.6976 | Glyma16g19540.1 | AT5G56180.1 | encodes a protein whose sequence is similar to actin-related proteins (ARPs) in other organisms. Member of nuclear ARP family of genes. |

| | | | | | |
|---|---|---|---|---|---|
| Gma.17273.1.A1_at | 0.05811 | 1.6967 | Glyma08g07820.1 | AT5G64170.2 | dentin sialophosphoprotein-related; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 17 plant structures; EXPRESSED DURING: 9 growth stages; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT3G54500.2); Has 117 Blast hits to 104 proteins in 33 species: Archae - 2; Bacteria - 18; Metazoa - 19; Fungi - 0; Plants - 57; Viruses - 0; Other Eukaryotes - 21 (source: NCBI BLink). |
| Gma.12481.1.S1_s_at | 0.0963 | 1.6917 | Glyma01g13170.2 | AT5G46800.1 | Seedling lethal mutation; Mitochondrial Carnitine Acyl Carrier-Like Protein |
| GmaAffx.68946.1.S1_at | 0.06925 | 1.6792 | Glyma08g25180.1 | AT1G79350.1 | embryo defective 1135 (EMB1135); FUNCTIONS IN: protein binding, DNA binding, zinc ion binding; INVOLVED IN: embryonic development ending in seed dormancy, regulation of transcription, DNA-dependent; LOCATED IN: chloroplast; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Zinc finger, RING/FYVE/PHD-type (InterPro:IPR013083), Zinc finger, PHD-type (InterPro:IPR001965), Zinc finger, LSD1-type (InterPro:IPR005735), Zinc finger, FYVE/PHD-type (InterPro:IPR011011); BEST Arabidopsis thaliana protein match is: ATXR6; DNA binding / protein binding (TAIR:AT5G24330.1); Has 3516 Blast hits to 3129 proteins in 248 species: Archae - 2; Bacteria - 434; Metazoa - 2253; Fungi - 273; Plants - 285; Viruses - 34; Other Eukaryotes - 235 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.38948.1.S1_at | 0.08725 | 1.6755 | Ambiguous Hit | AT3G22790.1 | kinase interacting family protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: plasma membrane; EXPRESSED IN: 20 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Prefoldin (InterPro:IPR009053), KIP1-like (InterPro:IPR011684); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT4G14760.1); Has 167995 Blast hits to 62112 proteins in 2147 species: Archae - 2214; Bacteria - 23421; Metazoa - 81973; Fungi - 11901; Plants - 6278; Viruses - 941; Other Eukaryotes - 41267 (source: NCBI BLink). |
| GmaAffx.52582.1.S1_at | 0.03169 | 1.6738 | Glyma03g35800.1 | AT5G22250.1 | CCR4-NOT transcription complex protein, putative; FUNCTIONS IN: ribonuclease activity, nucleic acid binding; INVOLVED IN: RNA modification; LOCATED IN: nucleus; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 11 growth stages; CONTAINS InterPro DOMAIN/s: Ribonuclease CAF1 (InterPro:IPR006941), Polynucleotidyl transferase, Ribonuclease H fold (InterPro:IPR012337); BEST Arabidopsis thaliana protein match is: CCR4-NOT transcription complex protein, putative (TAIR:AT3G44260.1); Has 630 Blast hits to 622 proteins in 157 species: Archae - 0; Bacteria - 0; Metazoa - 218; Fungi - 95; Plants - 219; Viruses - 0; Other Eukaryotes - 98 (source: NCBI BLink). |
| Gma.9160.1.A1_at | 0.03808 | 1.6686 | Glyma13g40850.4 | AT5G59800.1 | Protein containing methyl-CpG-binding domain.Has sequence similarity to human MBD proteins. Interacts with arginine methyltransferase 11. (AtPRMT11 |

| | | | | | |
|---|---|---|---|---|---|
| Gma.10888.2.A1_s_at | 0.04002 | 1.6674 | Glyma17g29800.2 | AT5G25560.1 | zinc finger (C3HC4-type RING finger) family protein; FUNCTIONS IN: protein binding, zinc ion binding; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Zinc finger, CTCHY-type (InterPro:IPR017921), Zinc finger, CHY-type (InterPro:IPR008913), Zinc finger, RING-type (InterPro:IPR001841), Zinc finger, RING/FYVE/PHD-type (InterPro:IPR013083); BEST Arabidopsis thaliana protein match is: zinc finger (C3HC4-type RING finger) family protein (TAIR:AT5G22920.1); Has 631 Blast hits to 629 proteins in 118 species: Archae - 2; Bacteria - 0; Metazoa - 274; Fungi - 89; Plants - 162; Viruses - 0; Other Eukaryotes - 104 (source: NCBI BLink). |
| GmaAffx.26500.1.S1_at | 0.04141 | 1.6636 | Glyma11g20130.1 | AT5G60760.1 | 2-phosphoglycerate kinase-related; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 12 plant structures; EXPRESSED DURING: 7 growth stages; BEST Arabidopsis thaliana protein match is: 2-phosphoglycerate kinase-related (TAIR:AT3G45090.1); Has 1178 Blast hits to 911 proteins in 165 species: Archae - 61; Bacteria - 75; Metazoa - 269; Fungi - 58; Plants - 88; Viruses - 45; Other Eukaryotes - 582 (source: NCBI BLink). |
| GmaAffx.27463.1.S1_at | 0.07583 | 1.6541 | Glyma11g00310.1 | AT5G02860.1 | pentatricopeptide (PPR) repeat-containing protein; INVOLVED IN: biological_process unknown; EXPRESSED IN: 18 plant structures; EXPRESSED DURING: 12 growth stages; CONTAINS InterPro DOMAIN/s: Pentatricopeptide repeat (InterPro:IPR002885); BEST Arabidopsis thaliana protein match is: pentatricopeptide (PPR) repeat-containing protein (TAIR:AT2G18940.1); Has 32008 Blast hits to 6661 proteins in 215 species: Archae - 4; Bacteria - 67; Metazoa - 1584; Fungi - 953; Plants - 27354; Viruses - 0; Other Eukaryotes - 2046 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.9087.1.A1_at | 0.05763 | 1.6405 | Glyma12g07010.2 | AT5G59950.4 | RNA and export factor-binding protein, putative; FUNCTIONS IN: nucleotide binding, nucleic acid binding; INVOLVED IN: biological_process unknown; LOCATED IN: chloroplast; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: RNA recognition motif, RNP-1 (InterPro:IPR000504), Nucleotide-binding, alpha-beta plait (InterPro:IPR012677); BEST Arabidopsis thaliana protein match is: RNA and export factor-binding protein, putative (TAIR:AT5G02530.1); Has 4318 Blast hits to 3983 proteins in 353 species: Archae - 2; Bacteria - 318; Metazoa - 2167; Fungi - 747; Plants - 681; Viruses - 0; Other Eukaryotes - 403 (source: NCBI BLink). |
| Gma.397.1.A1_s_at | 0.04141 | 1.635 | Glyma02g04620.1 | AT2G22500.1 | Encodes one of the mitochondrial dicarboxylate carriers (DIC): DIC1 (AT2G22500), DIC2 (AT4G24570), DIC3 (AT5G09470). |
| Gma.5783.2.S1_s_at | 0.06717 | 1.6322 | Glyma15g12210.1 | AT5G47550.1 | cysteine protease inhibitor, putative / cystatin, putative; FUNCTIONS IN: cysteine-type endopeptidase inhibitor activity; INVOLVED IN: biological_process unknown; LOCATED IN: cell wall; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Proteinase inhibitor I25, cystatin, conserved site (InterPro:IPR018073), Proteinase inhibitor I25, cystatin (InterPro:IPR000010); BEST Arabidopsis thaliana protein match is: cysteine protease inhibitor family protein / cystatin family protein (TAIR:AT4G16500.1); Has 507 Blast hits to 483 proteins in 88 species: Archae - 0; Bacteria - 2; Metazoa - 12; Fungi - 0; Plants - 482; Viruses - 0; Other Eukaryotes - 11 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.16649.1.S1_s_at | 0.04759 | 1.6303 | Glyma18g01550.1 | AT3G26100.2 | regulator of chromosome condensation (RCC1) family protein; FUNCTIONS IN: Ran GTPase binding; INVOLVED IN: biological_process unknown; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Regulator of chromosome condensation/beta-lactamase-inhibitor protein II (InterPro:IPR009091), Regulator of chromosome condensation, RCC1 (InterPro:IPR000408); BEST Arabidopsis thaliana protein match is: regulator of chromosome condensation (RCC1) family protein (TAIR:AT3G15430.2); Has 13018 Blast hits to 3813 proteins in 266 species: Archae - 62; Bacteria - 1274; Metazoa - 6169; Fungi - 625; Plants - 1234; Viruses - 2; Other Eukaryotes - 3652 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.88042.1.A1_at | 0.07994 | 1.6301 | Glyma05g08780.1 | AT5G14580.1 | polyribonucleotide nucleotidyltransferase, putative; FUNCTIONS IN: polyribonucleotide nucleotidyltransferase activity, 3'-5'-exoribonuclease activity, RNA binding, nucleic acid binding; INVOLVED IN: mRNA catabolic process, RNA processing; EXPRESSED IN: 16 plant structures; EXPRESSED DURING: 7 growth stages; CONTAINS InterPro DOMAIN/s: Nucleic acid-binding, OB-fold (InterPro:IPR012340), K Homology, type 1, subgroup (InterPro:IPR018111), Exoribonuclease, phosphorolytic domain 2 (InterPro:IPR015847), S1, RNA binding (InterPro:IPR003029), Polynucleotide phosphorylase, phosphorolytic RNA-binding, bacterial/organelle-type (InterPro:IPR015848), Nucleic acid-binding, OB-fold-like (InterPro:IPR016027), K Homology (InterPro:IPR004087), Exoribonuclease, phosphorolytic domain 1 (InterPro:IPR001247), Polyribonucleotide nucleotidyltransferase (InterPro:IPR012162); BEST Arabidopsis thaliana protein match is: RIF10 (resistant to inhibition with FSM 10); 3'-5'-exoribonuclease/ RNA binding / nucleic acid binding / polyribonucleotide nucleotidyltransferase (TAIR:AT3G03710.1); Has 19856 B |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.2689.1.S1_at | 0.09622 | 1.6207 | Glyma11g14580.1 | AT5G59550.1 | zinc finger (C3HC4-type RING finger) family protein; FUNCTIONS IN: protein binding, zinc ion binding; INVOLVED IN: response to chitin; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Zinc finger, RING-type (InterPro:IPR001841), Zinc finger, RING/FYVE/PHD-type (InterPro:IPR013083), Protein of unknown function DUF1117 (InterPro:IPR010543); BEST Arabidopsis thaliana protein match is: zinc finger (C3HC4-type RING finger) family protein (TAIR:AT3G46620.1); Has 6207 Blast hits to 6188 proteins in 212 species: Archae - 0; Bacteria - 6; Metazoa - 2070; Fungi - 537; Plants - 2564; Viruses - 41; Other Eukaryotes - 989 (source: NCBI BLink). |
| GmaAffx.40997.1.S1_at | 0.06802 | 1.5995 | Glyma04g15120.1 | AT5G04940.2 | Encodes a SU(VAR)3-9 homolog, a SET domain protein. Known SET domain proteins are involved in epigenetic control of gene expression and act as histone methyltransferases. There are 10 SUVH genes in Arabidopsis and members of this subfamily of the SET proteins have an additional conserved SRA domain. |
| Gma.1189.1.S1_at | 0.09965 | 1.5861 | Glyma20g37730.2 | AT5G27540.2 | Encodes a protein with similarity to GTPases that is localized to the mitochondrion. Involved in embryogenesis, pollen tube growth and required for mitochondrial development. |
| GmaAffx.63706.1.S1_s_at | 0.05105 | 1.5643 | Glyma05g01160.1 | AT1G74520.1 | Part of the AtHVA22a family. Protein expression is ABA- and stress-inducible. |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.77170.1.S1_at | 0.05138 | 1.5572 | Glyma10g26240.3 | AT2G34260.1 | transducin family protein / WD-40 repeat family protein; FUNCTIONS IN: nucleotide binding; LOCATED IN: CUL4 RING ubiquitin ligase complex, heterotrimeric G-protein complex; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: WD40 repeat-like (InterPro:IPR011046), WD40 repeat, region (InterPro:IPR017986), WD40/YVTN repeat-like (InterPro:IPR015943), WD40 repeat (InterPro:IPR001680), WD repeat protein 55 (InterPro:IPR017422); BEST Arabidopsis thaliana protein match is: transducin family protein / WD-40 repeat family protein (TAIR:AT2G43770.1); Has 16318 Blast hits to 9709 proteins in 465 species: Archae - 44; Bacteria - 3325; Metazoa - 6190; Fungi - 3093; Plants - 1110; Viruses - 0; Other Eukaryotes - 2556 (source: NCBI BLink). |
| GmaAffx.93010.1.S1_s_at | 0.07409 | 1.5558 | Glyma04g37370.3 | AT5G63490.1 | CBS domain-containing protein / octicosapeptide/Phox/Bem1 (PB1) domain-containing protein; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Octicosapeptide/Phox/Bem1p (InterPro:IPR000270), Cystathionine beta-synthase, core (InterPro:IPR000644); BEST Arabidopsis thaliana protein match is: CBS domain-containing protein / octicosapeptide/Phox/Bem1 (PB1) domain-containing protein (TAIR:AT5G50640.1); Has 5094 Blast hits to 4096 proteins in 849 species: Archae - 605; Bacteria - 3227; Metazoa - 2; Fungi - 80; Plants - 120; Viruses - 0; Other Eukaryotes - 1060 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.3661.1.S1_at | 0.02697 | 1.5537 | Glyma04g03860.1 | AT1G47710.1 | serpin, putative / serine protease inhibitor, putative; FUNCTIONS IN: serine-type endopeptidase inhibitor activity, cysteine-type endopeptidase inhibitor activity; INVOLVED IN: biological_process unknown; LOCATED IN: apoplast; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Protease inhibitor I4, serpin, plant (InterPro:IPR015554), Protease inhibitor I4, serpin (InterPro:IPR000215); BEST Arabidopsis thaliana protein match is: serpin, putative / serine protease inhibitor, putative (TAIR:AT3G45220.1); Has 4823 Blast hits to 4760 proteins in 347 species: Archae - 52; Bacteria - 220; Metazoa - 3736; Fungi - 1; Plants - 230; Viruses - 419; Other Eukaryotes - 165 (source: NCBI BLink). |
| GmaAffx.71637.1.S1_at | 0.05603 | 1.5446 | Glyma07g13730.2 | AT1G02140.1 | MAGO NASHI (MAGO); FUNCTIONS IN: protein binding; INVOLVED IN: embryonic development ending in seed dormancy, pollen tube guidance, sex determination; LOCATED IN: in 6 components; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Mago nashi protein (InterPro:IPR004023); Has 348 Blast hits to 348 proteins in 146 species: Archae - 0; Bacteria - 0; Metazoa - 176; Fungi - 64; Plants - 47; Viruses - 0; Other Eukaryotes - 61 (source: NCBI BLink). |
| Gma.11125.1.S1_at | 0.03261 | 1.5337 | Glyma10g21350.2 | AT3G57410.1 | Encodes a protein with high homology to animal villin. |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.22343.1.S1_at | 0.05809 | 1.5276 | Glyma15g28010.2 | AT1G54320.1 | LEM3 (ligand-effect modulator 3) family protein / CDC50 family protein; INVOLVED IN: biological_process unknown; LOCATED IN: chloroplast, membrane; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF284, transmembrane eukaryotic (InterPro:IPR005045); BEST Arabidopsis thaliana protein match is: ALIS1 (ALA-INTERACTING SUBUNIT 1); phospholipid transporter (TAIR:AT3G12740.1); Has 648 Blast hits to 646 proteins in 159 species: Archae - 0; Bacteria - 0; Metazoa - 218; Fungi - 139; Plants - 105; Viruses - 0; Other Eukaryotes - 186 (source: NCBI BLink). |
| Gma.11950.1.S1_at | 0.09495 | 1.5257 | No Soybean Match Identified, E<10E-30, Perecent ID>95% | AT3G07220.1 | transcriptional activator, putative; FUNCTIONS IN: transcription activator activity; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: guard cell; CONTAINS InterPro DOMAIN/s: SMAD/FHA domain (InterPro:IPR008984), Forkhead-associated (InterPro:IPR000253); BEST Arabidopsis thaliana protein match is: forkhead-associated domain-containing protein / FHA domain-containing protein (TAIR:AT3G07260.1); Has 336 Blast hits to 335 proteins in 110 species: Archae - 0; Bacteria - 0; Metazoa - 113; Fungi - 131; Plants - 59; Viruses - 0; Other Eukaryotes - 33 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.70698.1.A1_at | 0.07888 | 1.5229 | Glyma16g01390.4 | AT1G56220.3 | dormancy/auxin associated family protein; FUNCTIONS IN: molecular_function unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Dormancyauxin associated (InterPro:IPR008406); BEST Arabidopsis thaliana protein match is: dormancy/auxin associated protein-related (TAIR:AT1G54070.1); Has 130 Blast hits to 129 proteins in 26 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 129; Viruses - 0; Other Eukaryotes - 1 (source: NCBI BLink). |
| Gma.10202.1.A1_at | 0.07985 | 1.4692 | Glyma08g41670.1 | AT5G08560.2 | transducin family protein / WD-40 repeat family protein; FUNCTIONS IN: nucleotide binding; LOCATED IN: cellular_component unknown; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: WD40 repeat-like (InterPro:IPR011046), WD40 repeat, region (InterPro:IPR017986), WD40/YVTN repeat-like (InterPro:IPR015943), WD40 repeat (InterPro:IPR001680), CTLH, C-terminal to LisH motif (InterPro:IPR006595), LisH dimerisation motif (InterPro:IPR006594); BEST Arabidopsis thaliana protein match is: transducin family protein / WD-40 repeat family protein (TAIR:AT5G43920.1); Has 43626 Blast hits to 21368 proteins in 612 species: Archae - 68; Bacteria - 5676; Metazoa - 19918; Fungi - 8215; Plants - 3870; Viruses - 6; Other Eukaryotes - 5873 (source: NCBI BLink). |
| Gma.15650.1.A1_at | 0.08694 | 1.4546 | Glyma15g12080.1 | AT2G32300.1 | Encodes a uclacyanin, a protein precursor that is closely related to precursors of stellacyanins and a blue copper protein from pea pods. |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.1165.1.S1_at | 0.0963 | 1.4392 | Glyma04g09710.1 | AT5G49700.1 | DNA-binding protein-related; INVOLVED IN: biological_process unknown; EXPRESSED IN: shoot apex, embryo, inflorescence meristem; EXPRESSED DURING: D bilateral stage; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF296 (InterPro:IPR005175), Predicted AT-hook DNA-binding (InterPro:IPR014476); BEST Arabidopsis thaliana protein match is: DNA-binding protein-related (TAIR:AT1G14490.1); Has 424 Blast hits to 423 proteins in 20 species: Archae - 0; Bacteria - 2; Metazoa - 3; Fungi - 2; Plants - 417; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| Gma.13216.1.A1_at | 0.05172 | 1.4301 | Glyma02g47420.1 | AT1G07380.1 | ceramidase family protein; FUNCTIONS IN: ceramidase activity; INVOLVED IN: biological_process unknown; LOCATED IN: plasma membrane; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Neutral/alkaline nonlysosomal ceramidase (InterPro:IPR006823); BEST Arabidopsis thaliana protein match is: ceramidase family protein (TAIR:AT2G38010.1); Has 458 Blast hits to 375 proteins in 133 species: Archae - 0; Bacteria - 160; Metazoa - 157; Fungi - 68; Plants - 30; Viruses - 0; Other Eukaryotes - 43 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.1997.1.S1_at | 0.0687 | 1.4131 | Glyma04g02510.2 | AT1G20110.1 | zinc finger (FYVE type) family protein; FUNCTIONS IN: phosphoinositide binding; INVOLVED IN: signal transduction; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Zinc finger, FYVE-type (InterPro:IPR000306), Zinc finger, FYVE-related (InterPro:IPR017455), Zinc finger, RING/FYVE/PHD-type (InterPro:IPR013083), Zinc finger, FYVE/PHD-type (InterPro:IPR011011); BEST Arabidopsis thaliana protein match is: phosphatidylinositol-4-phosphate 5-kinase family protein (TAIR:AT3G14270.1); Has 25361 Blast hits to 16508 proteins in 676 species: Archae - 11; Bacteria - 1450; Metazoa - 10130; Fungi - 5149; Plants - 3412; Viruses - 575; Other Eukaryotes - 4634 (source: NCBI BLink). |
| Gma.854.1.A1_at | 0.09823 | 1.412 | Glyma11g01020.1 | AT3G22590.1 | RNA pol II accessory factor Cdc73 family protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: RNA polymerase II accessory factor, Cdc73 (InterPro:IPR007852); Has 392 Blast hits to 300 proteins in 144 species: Archae - 0; Bacteria - 0; Metazoa - 227; Fungi - 87; Plants - 21; Viruses - 0; Other Eukaryotes - 57 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.11803.1.A1_at | 0.05226 | 1.4105 | Glyma11g11240.1 | AT4G38710.1 | glycine-rich protein; FUNCTIONS IN: translation initiation factor activity; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Plant specific eukaryotic initiation factor 4B (InterPro:IPR010433); BEST Arabidopsis thaliana protein match is: EIF4B1; translation initiation factor (TAIR:AT3G26400.1); Has 13881 Blast hits to 7241 proteins in 523 species: Archae - 38; Bacteria - 1004; Metazoa - 5553; Fungi - 1110; Plants - 508; Viruses - 195; Other Eukaryotes - 5473 (source: NCBI BLink). |
| Gma.3789.2.A1_at | 0.07816 | 1.3751 | Glyma04g11560.1 | AT1G78010.1 | tRNA modification GTPase, putative; FUNCTIONS IN: GTP binding, GTPase activity; INVOLVED IN: small GTPase mediated signal transduction, tRNA modification; LOCATED IN: intracellular; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: GTP-binding (InterPro:IPR005289), Ras GTPase (InterPro:IPR001806), Small GTP-binding protein (InterPro:IPR005225), tRNA modification GTPase TrmE (InterPro:IPR004520), GTP-binding protein, HSR1-related (InterPro:IPR002917); BEST Arabidopsis thaliana protein match is: emb2738 (embryo defective 2738); GTP binding (TAIR:AT3G12080.2); Has 24592 Blast hits to 21257 proteins in 1767 species: Archae - 217; Bacteria - 15891; Metazoa - 407; Fungi - 184; Plants - 146; Viruses - 0; Other Eukaryotes - 7747 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.3847.1.S1_s_at | 0.03623 | 1.3719 | Glyma04g37370.3 | AT5G63490.1 | CBS domain-containing protein / octicosapeptide/Phox/Bem1 (PB1) domain-containing protein; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Octicosapeptide/Phox/Bem1p (InterPro:IPR000270), Cystathionine beta-synthase, core (InterPro:IPR000644); BEST Arabidopsis thaliana protein match is: CBS domain-containing protein / octicosapeptide/Phox/Bem1 (PB1) domain-containing protein (TAIR:AT5G50640.1); Has 5094 Blast hits to 4096 proteins in 849 species: Archae - 605; Bacteria - 3227; Metazoa - 2; Fungi - 80; Plants - 120; Viruses - 0; Other Eukaryotes - 1060 (source: NCBI BLink). |
| GmaAffx.230.2.S1_at | 0.05158 | 1.3616 | Glyma07g34330.1 | AT2G01460.1 | ATP binding / kinase; FUNCTIONS IN: kinase activity, ATP binding; INVOLVED IN: biosynthetic process; LOCATED IN: cellular_component unknown; EXPRESSED IN: 6 plant structures; EXPRESSED DURING: 4 anthesis; BEST Arabidopsis thaliana protein match is: phosphoribulokinase/uridine kinase family protein (TAIR:AT1G26190.1); Has 1426 Blast hits to 1171 proteins in 476 species: Archae - 7; Bacteria - 966; Metazoa - 112; Fungi - 108; Plants - 88; Viruses - 2; Other Eukaryotes - 143 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.2054.1.S1_at | 0.09251 | 1.3502 | No Soybean Match Identified, E<10E-30, Perecent ID>95% | AT3G47550.6 | zinc finger (C3HC4-type RING finger) family protein; FUNCTIONS IN: protein binding, zinc ion binding; INVOLVED IN: response to salt stress; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Zinc finger, RING/FYVE/PHD-type (InterPro:IPR013083), Zinc finger, RING-CH-type (InterPro:IPR011016); BEST Arabidopsis thaliana protein match is: zinc finger (C3HC4-type RING finger) family protein (TAIR:AT5G62460.1); Has 1346 Blast hits to 1345 proteins in 152 species: Archae - 0; Bacteria - 0; Metazoa - 706; Fungi - 79; Plants 322; Viruses - 29; Other Eukaryotes - 210 (source: NCBI BLink). |
| Gma.6698.1.A1_at | 0.06021 | 1.3501 | Glyma19g43970.1 | AT5G18410.2 | distorted trichomes and exhibits a diffuse actin cytoskeleton |
| Gma.400.1.S1_at | 0.08684 | 1.3461 | No Soybean Match Identified, E<10E-30, Perecent ID>95% | AT3G09350.2 | armadillo/beta-catenin repeat family protein; FUNCTIONS IN: binding; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Armadillo-like helical (InterPro:IPR011989), Armadillo-type fold (InterPro:IPR016024); BEST Arabidopsis thaliana protein match is: armadillo/beta-catenin repeat family protein (TAIR:AT3G53800.1); Has 426 Blast hits to 422 proteins in 144 species: Archae - 0; Bacteria - 0; Metazoa - 169; Fungi - 107; Plants - 93; Viruses - 0; Other Eukaryotes - 57 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.16938.1.S1_at | 0.07716 | 1.3117 | Glyma14g11780.1 | AT5G10840.1 | endomembrane protein 70, putative; LOCATED IN: integral to membrane, Golgi apparatus, plasma membrane, plant-type cell wall; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Nonaspanin (TM9SF) (InterPro:IPR004240); BEST Arabidopsis thaliana protein match is: endomembrane protein 70, putative (TAIR:AT5G25100.1); Has 1057 Blast hits to 1040 proteins in 178 species: Archae - 0; Bacteria - 38; Metazoa - 446; Fungi - 142; Plants - 235; Viruses - 0; Other Eukaryotes - 196 (source: NCBI BLink). |
| Gma.6223.1.S1_at | 0.05105 | 1.2986 | Glyma15g39100.1 | AT3G14610.1 | putative cytochrome P450 |
| Gma.4997.2.S1_at | 0.05719 | 1.2968 | Glyma08g02170.1 | AT4G11270.1 | transducin family protein / WD-40 repeat family protein; FUNCTIONS IN: nucleotide binding; INVOLVED IN: biological_process unknown; LOCATED IN: cytosol; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: WD40 repeat-like (InterPro:IPR011046), WD40 repeat, region (InterPro:IPR017986), WD40/YVTN repeat-like (InterPro:IPR015943), WD40 repeat (InterPro:IPR001680); Has 2796 Blast hits to 2122 proteins in 213 species: Archae - 0; Bacteria - 517; Metazoa - 1103; Fungi - 733; Plants - 103; Viruses - 0; Other Eukaryotes - 340 (source: NCBI BLink). |
| GmaAffx.39219.1.S1_at | 0.05811 | 1.272 | Glyma18g13250.1 | AT2G38020.1 | necessary for proper vacuole formation and morphogenesis in Arabidopsis |
| GmaAffx.24020.1.S1_at | 0.09963 | 1.242 | Glyma18g47410.1 | AT3G61570.1 | This gene is predicted to encode a protein that functions as a Golgi apparatus structural component known as a golgin in mammals and yeast. A fluorescently-tagged version of GC3 co-localizes with Golgi markers, and this localization appears to be replicated using the C-terminal (161 aa) portion of the protein. |

| | | | | | |
|---|---|---|---|---|---|
| Gma.10688.2.S1_at | 0.08725 | 1.2413 | Glyma02g44970.1 | AT3G57090.1 | Encodes a protein with similarity to yeast FIS proteins. These membrane anchored proteins bind DRP proteins and function during organelle division. FIS1B is expressed ubiquitously and appears to be involved in peroxisome division. |
| Gma.1538.1.S1_a_at | 0.05759 | 1.2067 | Glyma17g01800.1 | AT1G11650.2 | RBP45B; FUNCTIONS IN: RNA binding; EXPRESSED IN: 26 plant structures; EXPRESSED DURING: 16 growth stages; CONTAINS InterPro DOMAIN/s: RNA recognition motif, RNP-1 (InterPro:IPR000504), Nucleotide-binding, alpha-beta plait (InterPro:IPR012677); BEST Arabidopsis thaliana protein match is: ATRBP45C; RNA binding (TAIR:AT4G27000.1); Has 22877 Blast hits to 14020 proteins in 588 species: Archae - 14; Bacteria - 1484; Metazoa - 13266; Fungi - 2464; Plants - 3532; Viruses - 0; Other Eukaryotes - 2117 (source: NCBI BLink). |
| Gma.16713.1.S1_at | 0.09559 | 1.1914 | No Soybean Match Identified, E<10E-30, Perecent ID>95% | AT5G48970.1 | mitochondrial substrate carrier family protein; FUNCTIONS IN: binding; INVOLVED IN: transport, mitochondrial transport; LOCATED IN: mitochondrial inner membrane, membrane; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Mitochondrial carrier protein (InterPro:IPR002067), Mitochondrial substrate carrier (InterPro:IPR001993), Mitochondrial substrate/solute carrier (InterPro:IPR018108); BEST Arabidopsis thaliana protein match is: mitochondrial substrate carrier family protein (TAIR:AT3G21390.1); Has 21769 Blast hits to 10177 proteins in 363 species: Archae - 0; Bacteria - 0; Metazoa - 10956; Fungi - 6020; Plants - 2852; Viruses - 0; Other Eukaryotes - 1941 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.14088.1.A1_at | 0.05229 | -6.1646 | Glyma13g08190.1 | AT5G53190.1 | nodulin MtN3 family protein; INVOLVED IN: biological_process unknown; LOCATED IN: integral to membrane, membrane; EXPRESSED IN: 8 plant structures; EXPRESSED DURING: 4 anthesis, C globular stage, petal differentiation and expansion stage, E expanded cotyledon stage; CONTAINS InterPro DOMAIN/s: MtN3 and saliva related transmembrane protein, conserved region (InterPro:IPR018169), RAG1-activating protein 1 homologue (InterPro:IPR018179), MtN3 and saliva related transmembrane protein (InterPro:IPR004316); BEST Arabidopsis thaliana protein match is: nodulin MtN3 family protein (TAIR:AT1G21460.1); Has 523 Blast hits to 518 proteins in 92 species: Archae - 0; Bacteria - 0; Metazoa - 182; Fungi - 0; Plants - 291; Viruses - 0; Other Eukaryotes - 50 (source: NCBI BLink). |
| GmaAffx.34172.2.S1_at | 0.09965 | -5.5975 | Glyma04g41130.1 | AT1G09150.1 | pseudouridine synthase and archaeosine transglycosylase (PUA) domain-containing protein; FUNCTIONS IN: RNA binding; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: PUA-like (InterPro:IPR015947), PUA (InterPro:IPR002478), Translation machinery-associated RNA binding protein, predicted (InterPro:IPR016437), Uncharacterized domain 2 (InterPro:IPR004521); BEST Arabidopsis thaliana protein match is: eukaryotic translation initiation factor SUI1 family protein (TAIR:AT1G71350.1); Has 649 Blast hits to 647 proteins in 208 species: Archae - 99; Bacteria - 0; Metazoa - 263; Fungi - 102; Plants - 42; Viruses - 0; Other Eukaryotes - 143 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.9503.1.S1_at | 0.02316 | -5.2141 | Glyma18g24740.1 | AT5G65660.1 | hydroxyproline-rich glycoprotein family protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: plasma membrane; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; Has 31 Blast hits to 31 proteins in 11 species: Archae - 0; Bacteria - 0; Metazoa - 5; Fungi - 2; Plants - 21; Viruses - 0; Other Eukaryotes - 3 (source: NCBI BLink). |
| Gma.4716.2.S1_at | 0.05006 | -4.8694 | Glyma10g00220.1 | AT5G17540.1 | transferase family protein; FUNCTIONS IN: transferase activity, transferring acyl groups other than amino-acyl groups; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 9 plant structures; EXPRESSED DURING: 4 anthesis, F mature embryo stage, C globular stage, petal differentiation and expansion stage; CONTAINS InterPro DOMAIN/s: Transferase (InterPro:IPR003480); BEST Arabidopsis thaliana protein match is: CHAT (acetyl CoA:(Z)-3-hexen-1-ol acetyltransferase); acetyl CoA:(Z)-3-hexen-1-ol acetyltransferase (TAIR:AT3G03480.1); Has 1234 Blast hits to 1219 proteins in 99 species: Archae - 0; Bacteria - 2; Metazoa - 0; Fungi - 20; Plants - 1211; Viruses - 0; Other Eukaryotes - 1 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.51276.1.S1_at | 0.08156 | -4.3875 | Glyma05g32240.1 | AT1G49230.1 | zinc finger (C3HC4-type RING finger) family protein; FUNCTIONS IN: protein binding, zinc ion binding; EXPRESSED IN: 17 plant structures; EXPRESSED DURING: 11 growth stages; CONTAINS InterPro DOMAIN/s: Zinc finger, RING-type, conserved site (InterPro:IPR017907), Zinc finger, RING-type (InterPro:IPR001841), Zinc finger, RING/FYVE/PHD-type (InterPro:IPR013083); BEST Arabidopsis thaliana protein match is: zinc finger (C3HC4-type RING finger) family protein (TAIR:AT1G49200.1); Has 6176 Blast hits to 6158 proteins in 215 species: Archae - 0; Bacteria - 0; Metazoa - 2121; Fungi - 417; Plants - 2675; Viruses - 39; Other Eukaryotes - 924 (source: NCBI BLink). |
| Gma.6045.1.S1_at | 0.0481 | -4.0881 | Glyma02g39750.1 | AT2G30210.1 | putative laccase, a member of laccase family of genes (17 members in Arabidopsis). |
| GmaAffx.55366.1.S1_at | 0.09622 | -3.7991 | Glyma10g12060.1 | AT5G06900.1 | member of CYP93D |
| GmaAffx.69514.1.S1_at | 0.0472 | -3.7488 | Glyma11g18720.4 | AT2G27350.4 | OTU-like cysteine protease family protein; FUNCTIONS IN: cysteine-type peptidase activity; INVOLVED IN: biological_process unknown; LOCATED IN: chloroplast; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Ubiquitin-associated/translation elongation factor EF1B, N-terminal, eukaryote (InterPro:IPR015940), Ovarian tumour, otubain (InterPro:IPR003323); BEST Arabidopsis thaliana protein match is: SEC-C motif-containing protein / OTU-like cysteine protease family protein (TAIR:AT5G67170.2); Has 1538 Blast hits to 1134 proteins in 160 species: Archae - 0; Bacteria - 18; Metazoa - 694; Fungi - 186; Plants - 192; Viruses - 12; Other Eukaryotes - 436 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.54756.1.S1_at | 0.08672 | -3.7287 | Glyma17g33580.1 | AT2G13600.1 | pentatricopeptide (PPR) repeat-containing protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 7 plant structures; EXPRESSED DURING: F mature embryo stage, petal differentiation and expansion stage, D bilateral stage, E expanded cotyledon stage; CONTAINS InterPro DOMAIN/s: Pentatricopeptide repeat (InterPro:IPR002885); BEST Arabidopsis thaliana protein match is: pentatricopeptide (PPR) repeat-containing protein (TAIR:AT2G22070.1); Has 15471 Blast hits to 4950 proteins in 136 species: Archae - 0; Bacteria - 0; Metazoa - 72; Fungi - 40; Plants - 15040; Viruses - 0; Other Eukaryotes - 319 (source: NCBI BLink). |
| GmaAffx.74336.1.S1_at | 0.07553 | -3.5587 | Glyma13g40580.1 | AT3G45850.1 | kinesin motor protein-related; FUNCTIONS IN: microtubule motor activity, ATP binding; INVOLVED IN: microtubule-based movement; LOCATED IN: plasma membrane; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Kinesin, motor region (InterPro:IPR001752); BEST Arabidopsis thaliana protein match is: kinesin motor protein-related (TAIR:AT2G28620.1); Has 51195 Blast hits to 33804 proteins in 1398 species: Archae - 511; Bacteria - 4991; Metazoa - 25409; Fungi - 4042; Plants - 2382; Viruses - 131; Other Eukaryotes - 13729 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.61125.1.S1_at | 0.05004 | -3.4959 | Glyma08g43610.1 | AT5G55460.1 | protease inhibitor/seed storage/lipid transfer protein (LTP) family protein; FUNCTIONS IN: lipid binding; INVOLVED IN: lipid transport; LOCATED IN: endomembrane system; CONTAINS InterPro DOMAIN/s: Bifunctional inhibitor/plant lipid transfer protein/seed storage (InterPro:IPR016140), Plant lipid transfer protein/seed storage/trypsin-alpha amylase inhibitor (InterPro:IPR003612); BEST Arabidopsis thaliana protein match is: protease inhibitor/seed storage/lipid transfer protein (LTP) family protein (TAIR:AT5G55410.2); Has 71 Blast hits to 71 proteins in 13 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 71; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| Gma.13317.1.A1_at | 0.07755 | -3.4584 | Glyma02g31120.1 | AT3G57690.1 | Encodes a putative arabinogalactan-protein (AGP23). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.27805.1.A1_at | 0.02679 | -3.208 | Glyma11g36200.1 | AT2G22250.3 | ASPARTATE AMINOTRANSFERASE (AAT); FUNCTIONS IN: L-aspartate:2-oxoglutarate aminotransferase activity, transferase activity, transferring nitrogenous groups, pyridoxal phosphate binding, catalytic activity; INVOLVED IN: embryonic development ending in seed dormancy; LOCATED IN: chloroplast; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Aminotransferase, class I and II (InterPro:IPR004839), Pyridoxal phosphate-dependent transferase, major region (InterPro:IPR015424), Aminotransferases, class-I, pyridoxal-phosphate-binding site (InterPro:IPR004838), Pyridoxal phosphate-dependent transferase, major region, subdomain 1 (InterPro:IPR015421); BEST Arabidopsis thaliana protein match is: aminotransferase class I and II family protein (TAIR:AT1G77670.1); Has 32033 Blast hits to 32031 proteins in 1793 species: Archae - 684; Bacteria - 17916; Metazoa - 676; Fungi - 572; Plants - 943; Viruses - 0; Other Eukaryotes - 11242 (source: NCBI BLink). |
| GmaAffx.64032.1.S1_at | 0.03707 | -2.9282 | Glyma07g13450.2 | AT4G17670.1 | senescence-associated protein-related; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF581 (InterPro:IPR007650); BEST Arabidopsis thaliana protein match is: senescence-associated protein-related (TAIR:AT5G47060.1); Has 285 Blast hits to 285 proteins in 14 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 285; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.10521.1.A1_at | 0.06013 | -2.7417 | Glyma05g27970.1 | AT1G13710.1 | Encodes the cytochrome P450 CYP78A5 monooxygenase. Contributes to the generation of a growth-stimulating signal distinct from the classical phytohormones that prevents proliferation arrest, promoting organ growth. |
| GmaAffx.54969.1.S1_at | 0.06238 | -2.6903 | Ambiguous Hit | AT2G16700.1 | Encodes actin depolymerizing factor 5 (ADF5). |
| GmaAffx.50341.1.S1_at | 0.08973 | -2.6521 | Glyma06g19770.1 | AT1G24130.1 | transducin family protein / WD-40 repeat family protein; FUNCTIONS IN: nucleotide binding; INVOLVED IN: biological_process unknown; LOCATED IN: chloroplast, heterotrimeric G-protein complex; CONTAINS InterPro DOMAIN/s: WD40 repeat-like (InterPro:IPR011046), WD40 repeat, region (InterPro:IPR017986), WD40/YVTN repeat-like (InterPro:IPR015943), WD40 repeat (InterPro:IPR001680); BEST Arabidopsis thaliana protein match is: transducin family protein / WD-40 repeat family protein (TAIR:AT5G50120.1); Has 33915 Blast hits to 16327 proteins in 570 species: Archae - 40; Bacteria - 4362; Metazoa - 14549; Fungi - 6756; Plants - 3137; Viruses - 0; Other Eukaryotes - 5071 (source: NCBI BLink). |
| GmaAffx.91050.1.S1_at | 0.05105 | -2.6328 | Glyma20g02310.1 | AT1G64950.1 | member of CYP89A |
| Gma.8778.1.S1_at | 0.07119 | -2.6023 | Glyma13g18930.1 | AT4G28660.1 | Similar to PsbW subunit of photosystem II. |
| GmaAffx.55917.2.A1_at | 0.09381 | -2.5941 | Glyma18g45070.1 | AT5G24910.1 | member of CYP714A |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.57642.1.S1_at | 0.06018 | -2.4816 | Glyma18g48200.1 | AT3G25480.1 | rhodanese-like domain-containing protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: chloroplast thylakoid membrane, chloroplast; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Rhodanese-like (InterPro:IPR001763); BEST Arabidopsis thaliana protein match is: hydroxyproline-rich glycoprotein family protein (TAIR:AT4G01050.1); Has 48 Blast hits to 48 proteins in 10 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 48; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| GmaAffx.25263.1.S1_at | 0.05489 | -2.3746 | Glyma05g27980.1 | AT1G67360.1 | rubber elongation factor (REF) family protein; INVOLVED IN: biological_process unknown; LOCATED IN: vacuole; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Rubber elongation factor (InterPro:IPR008802); BEST Arabidopsis thaliana protein match is: rubber elongation factor (REF) protein-related (TAIR:AT2G47780.1); Has 76 Blast hits to 76 proteins in 14 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 76; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.16213.1.S1_at | 0.03154 | -2.3377 | Glyma20g12150.1 | AT5G21105.1 | L-ascorbate oxidase/ copper ion binding / oxidoreductase; FUNCTIONS IN: oxidoreductase activity, L-ascorbate oxidase activity, copper ion binding; INVOLVED IN: oxidation reduction; LOCATED IN: cell wall; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Cupredoxin (InterPro:IPR008972), Multicopper oxidase, type 2 (InterPro:IPR011706), Multicopper oxidase, type 3 (InterPro:IPR011707), Multicopper oxidase, copper-binding site (InterPro:IPR002355), L-ascorbate oxidase, plants (InterPro:IPR017760), Multicopper oxidase, type 1 (InterPro:IPR001117); BEST Arabidopsis thaliana protein match is: L-ascorbate oxidase, putative (TAIR:AT5G21100.1); Has 6896 Blast hits to 6267 proteins in 1013 species: Archae - 39; Bacteria - 2676; Metazoa - 535; Fungi - 2466; Plants - 798; Viruses - 0; Other Eukaryotes - 382 (source: NCBI BLink). |
| GmaAffx.81982.2.S1_at | 0.08139 | -2.2213 | Glyma10g19710.2 | AT3G56590.1 | hydroxyproline-rich glycoprotein family protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; BEST Arabidopsis thaliana protein match is: zinc finger (C3HC4-type RING finger) family protein (TAIR:AT3G10810.1); Has 22948 Blast hits to 13764 proteins in 833 species: Archae - 42; Bacteria - 2594; Metazoa - 8521; Fungi - 3551; Plants - 4342; Viruses - 788; Other Eukaryotes - 3110 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.5961.1.A1_at | 0.09706 | -2.2195 | Glyma14g08030.1 | AT1G75260.1 | oxidoreductase, acting on NADH or NADPH; FUNCTIONS IN: oxidoreductase activity, acting on NADH or NADPH; LOCATED IN: chloroplast; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT2G22795.1); Has 9000 Blast hits to 6650 proteins in 578 species: Archae - 17; Bacteria - 793; Metazoa - 3741; Fungi - 1063; Plants - 304; Viruses - 64; Other Eukaryotes - 3018 (source: NCBI BLink). |
| GmaAffx.76840.1.S1_at | 0.08755 | -2.1919 | Glyma15g14970.1 | AT3G06610.1 | DNA-binding enhancer protein-related; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; Has 145 Blast hits to 145 proteins in 71 species: Archae - 0; Bacteria - 0; Metazoa - 98; Fungi - 15; Plants - 20; Viruses - 0; Other Eukaryotes - 12 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.13939.1.S1_at | 0.0398 | -2.1611 | Glyma17g10420.1 | AT5G62670.1 | Arabidopsis H(+)-ATPase 11 (AHA11); FUNCTIONS IN: ATPase activity; INVOLVED IN: cation transport, metabolic process, ATP biosynthetic process; LOCATED IN: plasma membrane, membrane; EXPRESSED IN: 26 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: ATPase, P-type, ATPase-associated region (InterPro:IPR008250), ATPase, P-type cation-transporter, N-terminal (InterPro:IPR004014), Haloacid dehalogenase-like hydrolase (InterPro:IPR005834), ATPase, P-type, H+ transporting proton pump (InterPro:IPR000695), ATPase, P-type, K/Mg/Cd/Cu/Zn/Na/Ca/Na/H-transporter (InterPro:IPR001757), ATPase, P-type, plasma-membrane proton-efflux (InterPro:IPR006534), ATPase, P-type phosphorylation site (InterPro:IPR018303); BEST Arabidopsis thaliana protein match is: AHA4; ATPase/ hydrogen-exporting ATPase, phosphorylative mechanism (TAIR:AT3G47950.1); Has 20808 Blast hits to 18623 proteins in 1829 species: Archae - 405; Bacteria - 11444; Metazoa - 3332; Fungi - 1581; Plants - 1110; Viruses - 3; Other Eukaryotes - 2933 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.5994.1.S1_at | 0.05952 | -2.1491 | Glyma07g39590.1 | AT5G47550.1 | cysteine protease inhibitor, putative / cystatin, putative; FUNCTIONS IN: cysteine-type endopeptidase inhibitor activity; INVOLVED IN: biological_process unknown; LOCATED IN: cell wall; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Proteinase inhibitor I25, cystatin, conserved site (InterPro:IPR018073), Proteinase inhibitor I25, cystatin (InterPro:IPR000010); BEST Arabidopsis thaliana protein match is: cysteine protease inhibitor family protein / cystatin family protein (TAIR:AT4G16500.1); Has 507 Blast hits to 483 proteins in 88 species: Archae - 0; Bacteria - 2; Metazoa - 12; Fungi - 0; Plants - 482; Viruses - 0; Other Eukaryotes - 11 (source: NCBI BLink). |
| Gma.11635.1.S1_at | 0.08105 | -2.0701 | Glyma12g28760.1 | AT5G49920.1 | octicosapeptide/Phox/Bem1p (PB1) domain-containing protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; CONTAINS InterPro DOMAIN/s: Octicosapeptide/Phox/Bem1p (InterPro:IPR000270); BEST Arabidopsis thaliana protein match is: protein kinase family protein (TAIR:AT3G46920.1); Has 185 Blast hits to 185 proteins in 9 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 2; Plants - 183; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| GmaAffx.74241.2.A1_at | 0.03987 | -2.0478 | Glyma09g20270.2 | AT1G75130.1 | member of CYP721A |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.83465.1.A1_at | 0.05158 | -2.0148 | Glyma05g35730.1 | AT5G19670.1 | exostosin family protein; FUNCTIONS IN: catalytic activity; INVOLVED IN: biological_process unknown; LOCATED IN: membrane; EXPRESSED IN: 10 plant structures; EXPRESSED DURING: 6 growth stages; CONTAINS InterPro DOMAIN/s: Exostosin-like (InterPro:IPR004263); BEST Arabidopsis thaliana protein match is: exostosin family protein (TAIR:AT5G25820.1); Has 802 Blast hits to 800 proteins in 88 species: Archae - 0; Bacteria - 10; Metazoa - 236; Fungi - 4; Plants - 470; Viruses - 0; Other Eukaryotes - 82 (source: NCBI BLink). |
| GmaAffx.79307.1.S1_at | 0.03473 | -1.9778 | Glyma07g30720.1 | AT1G55890.1 | pentatricopeptide (PPR) repeat-containing protein; INVOLVED IN: biological_process unknown; LOCATED IN: mitochondrion; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Pentatricopeptide repeat (InterPro:IPR002885); BEST Arabidopsis thaliana protein match is: pentatricopeptide (PPR) repeat-containing protein (TAIR:AT3G13160.1); Has 17815 Blast hits to 5592 proteins in 168 species: Archae - 5; Bacteria - 22; Metazoa - 302; Fungi - 235; Plants - 16660; Viruses - 0; Other Eukaryotes - 591 (source: NCBI BLink). |
| Gma.17586.1.S1_at | 0.04002 | -1.9745 | Glyma18g14460.1 | AT3G10260.3 | reticulon family protein; LOCATED IN: mitochondrion, endoplasmic reticulum; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Reticulon (InterPro:IPR003388); BEST Arabidopsis thaliana protein match is: reticulon family protein (RTNLB6) (TAIR:AT3G61560.1); Has 821 Blast hits to 821 proteins in 61 species: Archae - 0; Bacteria - 0; Metazoa - 561; Fungi - 0; Plants - 251; Viruses - 0; Other Eukaryotes - 9 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.73771.1.S1_at | 0.06847 | -1.9474 | Glyma09g33700.1 | AT5G46790.1 | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 12 growth stages; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT4G17870.1); Has 179 Blast hits to 179 proteins in 19 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 178; Viruses - 0; Other Eukaryotes - 1 (source: NCBI BLink). |
| Gma.5449.1.S1_at | 0.05376 | -1.9472 | Glyma02g15910.2 | AT2G30620.1 | histone H1.2; FUNCTIONS IN: DNA binding; INVOLVED IN: nucleosome assembly; LOCATED IN: nucleolus, nucleus, nucleosome; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Winged helix repressor DNA-binding (InterPro:IPR011991), Histone H5 (InterPro:IPR005819), Histone H1/H5 (InterPro:IPR005818); BEST Arabidopsis thaliana protein match is: histone H1, putative (TAIR:AT1G06760.1); Has 17371 Blast hits to 9398 proteins in 1065 species: Archae - 39; Bacteria - 6415; Metazoa - 4684; Fungi - 1032; Plants - 746; Viruses - 137; Other Eukaryotes - 4318 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.77200.1.S1_at | 0.07658 | -1.9176 | Glyma08g12380.1 | AT2G34200.1 | zinc finger (C3HC4-type RING finger) family protein; FUNCTIONS IN: protein binding, zinc ion binding; LOCATED IN: membrane; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 12 growth stages; CONTAINS InterPro DOMAIN/s: Zinc finger, RING/FYVE/PHD-type (InterPro:IPR013083), Zinc finger, RING-CH-type (InterPro:IPR011016); BEST Arabidopsis thaliana protein match is: zinc finger (C3HC4-type RING finger) family protein (TAIR:AT5G05830.1); Has 286 Blast hits to 286 proteins in 47 species: Archae - 0; Bacteria - 0; Metazoa - 96; Fungi - 0; Plants - 187; Viruses - 0; Other Eukaryotes - 3 (source: NCBI BLink). |
| Gma.3869.1.S1_at | 0.09403 | -1.8595 | Glyma14g08230.2 | AT1G19380.1 | unknown protein; INVOLVED IN: biological_process unknown; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF1195 (InterPro:IPR010608); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT5G65650.1); Has 43 Blast hits to 43 proteins in 11 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 43; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| GmaAffx.15207.1.A1_at | 0.09351 | -1.8411 | Glyma11g35270.1 | AT2G40080.1 | Encodes a novel nuclear 111 amino-acid phytochrome-regulated component of a negative feedback loop involving the circadian clock central oscillator components CCA1 and LHY. ELF4 is necessary for light-induced expression of both CCA1 and LHY, and conversely, CCA1 and LHY act negatively on light-induced ELF4 expression. ELF4 promotes clock accuracy and is required for sustained rhythms in the absence of daily light/dark cycles. It is involved in the phyB-mediated constant red light induced seedling de-etiolation process and may function to coregulate the expression of a subset of phyB-regulated genes. |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.41544.1.S1_at | 0.09381 | -1.7587 | Glyma03g22450.1 | AT5G48460.1 | fimbrin-like protein, putative; FUNCTIONS IN: actin binding; INVOLVED IN: biological_process unknown; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Actinin-type, actin-binding, conserved site (InterPro:IPR001589), Calponin-homology (InterPro:IPR016146), Calponin-like actin-binding (InterPro:IPR001715); BEST Arabidopsis thaliana protein match is: FIM2 (FIMBRIN-LIKE PROTEIN 2); actin binding (TAIR:AT5G35700.1); Has 2573 Blast hits to 1976 proteins in 141 species: Archae - 0; Bacteria - 0; Metazoa - 2069; Fungi - 216; Plants - 75; Viruses - 0; Other Eukaryotes - 213 (source: NCBI BLink). |
| Gma.6956.1.A1_at | 0.06877 | -1.755 | Glyma04g37810.2 | AT5G07290.1 | AML4 A member of mei2-like gene family, predominantly plant-based family of genes encoding RNA binding proteins with characteristic presence of a highly conserved RNA binding motif first described in the mei2 gene of the fission yeast S. pombe. In silico analyses reveal nine mei2 -like genes in A. thaliana. They were grouped into four distinct clades, based on overall sequence similarity and subfamily-specific sequence elements. AML4 is a member of two sister clades of mei2-like gene family, AML1 through AML5, and belongs to the clade named ALM14. AML4 is expressed during embryo development (heart and torpedo stage) and in vegetative and floral apices. |

| | | | | | |
|---|---|---|---|---|---|
| Gma.5825.1.S1_x_at | 0.09586 | -1.7335 | Glyma13g17220.1 | AT1G74940.1 | senescence-associated protein-related; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF581 (InterPro:IPR007650); BEST Arabidopsis thaliana protein match is: senescence-associated protein-related (TAIR:AT1G19200.1); Has 261 Blast hits to 261 proteins in 15 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 261; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| Gma.8408.1.A1_at | 0.09193 | -1.7152 | Glyma16g03490.1 | AT1G14620.1 | DECOY (DECOY); FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; Has 181 Blast hits to 181 proteins in 92 species: Archae - 0; Bacteria - 2; Metazoa - 71; Fungi - 79; Plants - 19; Viruses - 0; Other Eukaryotes - 10 (source: NCBI BLink). |
| Gma.13547.1.A1_at | 0.04364 | -1.6824 | Glyma05g37810.2 | AT4G11440.1 | binding; FUNCTIONS IN: binding; INVOLVED IN: transport, mitochondrial transport; LOCATED IN: mitochondrial inner membrane, membrane; EXPRESSED IN: shoot apex, embryo, seed; EXPRESSED DURING: D bilateral stage, E expanded cotyledon stage; CONTAINS InterPro DOMAIN/s: Mitochondrial substrate carrier (InterPro:IPR001993), Mitochondrial substrate/solute carrier (InterPro:IPR018108); BEST Arabidopsis thaliana protein match is: SAMC1 (S-ADENOSYLMETHIONINE CARRIER 1); S-adenosylmethionine transmembrane transporter/ binding (TAIR:AT4G39460.1); Has 19661 Blast hits to 10258 proteins in 352 species: Archae - 0; Bacteria - 0; Metazoa - 9757; Fungi - 5517; Plants - 2554; Viruses - 0; Other Eukaryotes - 1833 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.57031.1.A1_at | 0.07755 | -1.6731 | Glyma11g06650.2 | AT4G36860.1 | zinc ion binding; FUNCTIONS IN: zinc ion binding; LOCATED IN: plasma membrane; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Zinc finger, LIM-type (InterPro:IPR001781), Ubiquitin interacting motif (InterPro:IPR003903); BEST Arabidopsis thaliana protein match is: DA1 (DA 1); ubiquitin binding (TAIR:AT1G19270.1); Has 1546 Blast hits to 1204 proteins in 114 species: Archae - 0; Bacteria - 2; Metazoa - 1348; Fungi - 12; Plants - 88; Viruses - 0; Other Eukaryotes - 96 (source: NCBI BLink). |
| Gma.1927.1.A1_at | 0.03473 | -1.6622 | Glyma01g37490.1 | AT4G36400.2 | Encodes a (D)-2-hydroxyglutarate dehydrogenase. |
| Gma.3888.3.S1_a_at | 0.07428 | -1.6585 | Glyma08g21300.1 | AT1G14520.1 | Encodes MIOX1. Belongs to myo-inositol oxygenase gene family. |
| GmaAffx.4773.1.A1_at | 0.05105 | -1.6573 | Glyma11g04610.1 | AT5G46640.1 | DNA-binding family protein; FUNCTIONS IN: DNA binding; EXPRESSED IN: 19 plant structures; EXPRESSED DURING: 6 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF296 (InterPro:IPR005175), AT hook, DNA-binding, conserved site (InterPro:IPR017956); BEST Arabidopsis thaliana protein match is: DNA-binding family protein (TAIR:AT4G17950.1); Has 568 Blast hits to 566 proteins in 77 species: Archae - 0; Bacteria - 71; Metazoa - 61; Fungi - 2; Plants - 415; Viruses - 0; Other Eukaryotes - 19 (source: NCBI BLink). |
| Gma.13387.1.A1_at | 0.07476 | -1.6542 | No Soybean Match Identified, E<10E-30, Perecent ID>95% | AT5G40660.1 | ATP12 protein-related; FUNCTIONS IN: protein binding; INVOLVED IN: proton-transporting ATP synthase complex assembly; LOCATED IN: cellular_component unknown; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: ATP12, ATPase F1F0-assembly protein (InterPro:IPR011419); Has 598 Blast hits to 598 proteins in 229 species: Archae - 0; Bacteria - 222; Metazoa - 101; Fungi - 70; Plants - 22; Viruses - 0; Other Eukaryotes - 183 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.1856.1.A1_at | 0.047 | -1.6462 | Glyma06g15760.1 | AT5G64580.1 | AAA-type ATPase family protein; FUNCTIONS IN: nucleoside-triphosphatase activity, metalloendopeptidase activity, nucleotide binding, zinc ion binding, ATP binding; LOCATED IN: chloroplast, chloroplast envelope; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: ATPase, AAA+ type, core (InterPro:IPR003593), ATPase, AAA-type, core (InterPro:IPR003959), ATPase, AAA-type, conserved site (InterPro:IPR003960), Peptidase M41, FtsH extracellular (InterPro:IPR011546); BEST Arabidopsis thaliana protein match is: VAR1 (VARIEGATED 1); ATP-dependent peptidase/ ATPase/ metallopeptidase (TAIR:AT5G42270.1); Has 29440 Blast hits to 27884 proteins in 1904 species: Archae - 945; Bacteria - 10493; Metazoa - 4130; Fungi - 2527; Plants - 1734; Viruses - 43; Other Eukaryotes - 9568 (source: NCBI BLink). |
| GmaAffx.5391.1.A1_at | 0.06877 | -1.6329 | Glyma07g08390.1 | AT4G23040.1 | UBX domain-containing protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cytosol; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: UBX (InterPro:IPR001012), UBA-like (InterPro:IPR009060); BEST Arabidopsis thaliana protein match is: SAY1 (TAIR:AT4G11740.1); Has 2849 Blast hits to 2369 proteins in 288 species: Archae - 2; Bacteria - 267; Metazoa - 1170; Fungi - 422; Plants - 198; Viruses - 20; Other Eukaryotes - 770 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.67477.1.S1_at | 0.06717 | -1.6171 | Glyma18g48620.1 | AT4G16830.1 | nuclear RNA-binding protein (RGGA); FUNCTIONS IN: RNA binding; INVOLVED IN: biological_process unknown; LOCATED IN: nucleus, cytoplasm; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Hyaluronan/mRNA binding protein (InterPro:IPR006861); BEST Arabidopsis thaliana protein match is: nuclear RNA-binding protein, putative (TAIR:AT5G47210.1); Has 21311 Blast hits to 10623 proteins in 811 species: Archae - 12; Bacteria - 6397; Metazoa - 7146; Fungi - 1662; Plants - 3005; Viruses - 285; Other Eukaryotes - 2804 (source: NCBI BLink). |
| Gma.8226.1.S1_at | 0.0721 | -1.6066 | Glyma08g45140.1 | AT5G58960.3 | Mutant plants display impaired light-regulation of the hypocotyl randomization response. |
| Gma.12148.1.S1_at | 0.07119 | -1.5995 | Glyma15g10750.1 | AT4G21860.1 | methionine sulfoxide reductase B 2 (MSRB2); FUNCTIONS IN: peptide-methionine-(S)-S-oxide reductase activity; INVOLVED IN: biological_process unknown; LOCATED IN: chloroplast, chloroplast stroma; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Methionine sulphoxide reductase B (InterPro:IPR002579), Mss4-like (InterPro:IPR011057); BEST Arabidopsis thaliana protein match is: methionine sulfoxide reductase domain-containing protein / SeIR domain-containing protein (TAIR:AT4G04800.1); Has 5933 Blast hits to 5932 proteins in 1231 species: Archae - 53; Bacteria - 2689; Metazoa - 222; Fungi - 85; Plants - 108; Viruses - 1; Other Eukaryotes - 2775 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.1010.1.S1_at | 0.07755 | -1.5922 | Glyma06g17900.2 | AT2G27110.3 | FAR1-related sequence 3 (FRS3); FUNCTIONS IN: zinc ion binding; INVOLVED IN: response to red or far red light; LOCATED IN: cellular_component unknown; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: MULE transposase, conserved domain (InterPro:IPR018289), Zinc finger, PMZ-type (InterPro:IPR006564), Zinc finger, SWIM-type (InterPro:IPR007527); BEST Arabidopsis thaliana protein match is: FRS5 (FAR1-related sequence 5); zinc ion binding (TAIR:AT4G38180.1); Has 641 Blast hits to 619 proteins in 25 species: Archae - 0; Bacteria - 0; Metazoa - 1; Fungi - 30; Plants - 610; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| Gma.10707.4.S1_at | 0.09965 | -1.5892 | Glyma13g40900.1 | AT5G02560.1 | Encodes HTA12, a histone H2A protein. |
| Gma.7311.1.A1_at | 0.05957 | -1.5863 | Glyma06g12210.1 | AT1G21690.1 | embryo defective 1968 (emb1968); FUNCTIONS IN: in 6 functions; INVOLVED IN: embryonic development ending in seed dormancy; LOCATED IN: DNA replication factor C complex, nucleolus; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: ATPase, AAA+ type, core (InterPro:IPR003593), ATPase, AAA-type, core (InterPro:IPR003959), Replication factor C (InterPro:IPR013748), DNA polymerase III clamp loader subunit, C-terminal (InterPro:IPR008921); BEST Arabidopsis thaliana protein match is: replication factor C 40 kDa, putative (TAIR:AT1G63160.1); Has 12080 Blast hits to 12043 proteins in 1582 species: Archae - 427; Bacteria - 5194; Metazoa - 735; Fungi - 600; Plants - 188; Viruses - 64; Other Eukaryotes - 4872 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.6522.1.A1_at | 0.05394 | -1.5849 | Glyma19g35690.1 | AT3G53650.1 | histone H2B, putative; FUNCTIONS IN: DNA binding; INVOLVED IN: nucleosome assembly; LOCATED IN: nucleus, nucleosome; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 12 growth stages; CONTAINS InterPro DOMAIN/s: Histone H2B (InterPro:IPR000558), Histone-fold (InterPro:IPR009072), Histone core (InterPro:IPR007125); BEST Arabidopsis thaliana protein match is: histone H2B, putative (TAIR:AT5G02570.1); Has 2705 Blast hits to 2699 proteins in 275 species: Archae - 0; Bacteria - 2; Metazoa - 1852; Fungi - 166; Plants - 360; Viruses - 0; Other Eukaryotes - 325 (source: NCBI BLink). |
| GmaAffx.4765.1.S1_at | 0.05957 | -1.5679 | Glyma16g05050.1 | AT3G02060.2 | DEAD/DEAH box helicase, putative; FUNCTIONS IN: transcription factor activity, helicase activity, ATP binding, ATP-dependent helicase activity, nucleic acid binding; INVOLVED IN: regulation of transcription, DNA-dependent; LOCATED IN: chloroplast; EXPRESSED IN: 16 plant structures; EXPRESSED DURING: 12 growth stages; CONTAINS InterPro DOMAIN/s: DNA/RNA helicase, DEAD/DEAH box type, N-terminal (InterPro:IPR011545), Transcription factor CarD (InterPro:IPR003711), DEAD-like helicase, N-terminal (InterPro:IPR014001), DNA/RNA helicase, C-terminal (InterPro:IPR001650), Helicase, superfamily 1 and 2, ATP-binding (InterPro:IPR014021); BEST Arabidopsis thaliana protein match is: ATP binding / ATP-dependent DNA helicase/ ATP-dependent helicase/ helicase/ nucleic acid binding (TAIR:AT2G01440.1); Has 20526 Blast hits to 20397 proteins in 1616 species: Archae - 162; Bacteria - 10276; Metazoa - 1293; Fungi - 732; Plants - 354; Viruses - 10; Other Eukaryotes - 7699 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.57844.1.S1_at | 0.04164 | -1.5305 | Glyma05g03320.1 | AT5G50400.1 | PURPLE ACID PHOSPHATASE 27 (PAP27); FUNCTIONS IN: protein serine/threonine phosphatase activity, acid phosphatase activity; INVOLVED IN: biological_process unknown; LOCATED IN: endomembrane system; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Purple acid phosphatase, N-terminal (InterPro:IPR015914), Metallophosphoesterase (InterPro:IPR004843), Purple acid phosphatase-like, N-terminal (InterPro:IPR008963); BEST Arabidopsis thaliana protein match is: PAP24 (PURPLE ACID PHOSPHATASE 24); acid phosphatase/ protein serine/threonine phosphatase (TAIR:AT4G24890.1); Has 985 Blast hits to 977 proteins in 173 species: Archae - 0; Bacteria - 156; Metazoa - 174; Fungi - 58; Plants - 412; Viruses - 0; Other Eukaryotes - 185 (source: NCBI BLink). |
| Gma.13764.1.A1_at | 0.09397 | -1.4955 | Glyma18g01670.1 | AT1G13030.1 | sphere organelles protein-related; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 19 plant structures; EXPRESSED DURING: 12 growth stages; Has 13894 Blast hits to 8686 proteins in 584 species: Archae - 12; Bacteria - 813; Metazoa - 5398; Fungi - 1230; Plants - 486; Viruses - 98; Other Eukaryotes - 5857 (source: NCBI BLink). |
| GmaAffx.20345.1.S1_at | 0.0917 | -1.472 | Glyma06g17990.1 | AT5G50375.1 | Converts pentacyclic cyclopropyl sterols to conventional tetracyclic sterols. CPI1 function during and just after division and support gravitropism by establishing polar PIN2 localization. Required for endocytosis of PIN2 |

| | | | | | |
|---|---|---|---|---|---|
| Gma.10441.1.S1_s_at | 0.09162 | -1.4616 | Glyma17g32570.1 | AT5G11090.1 | serine-rich protein-related; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; BEST Arabidopsis thaliana protein match is: serine-rich protein-related (TAIR:AT5G25280.2); Has 1617 Blast hits to 241 proteins in 57 species: Archae - 0; Bacteria - 4; Metazoa - 631; Fungi - 53; Plants - 88; Viruses - 0; Other Eukaryotes - 841 (source: NCBI BLink). |
| GmaAffx.87428.1.S1_at | 0.06925 | -1.4497 | Glyma19g42260.1 | AT5G27650.1 | PWWP domain-containing protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 12 growth stages; CONTAINS InterPro DOMAIN/s: PWWP (InterPro:IPR000313); BEST Arabidopsis thaliana protein match is: PWWP domain-containing protein (TAIR:AT3G05430.1); Has 6970 Blast hits to 4442 proteins in 299 species: Archae - 10; Bacteria - 363; Metazoa - 3592; Fungi - 627; Plants - 302; Viruses - 141; Other Eukaryotes - 1935 (source: NCBI BLink). |
| Gma.7657.1.S1_at | 0.05948 | -1.4429 | Glyma02g22730.2 | AT3G56590.2 | hydroxyproline-rich glycoprotein family protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; BEST Arabidopsis thaliana protein match is: zinc finger (C3HC4-type RING finger) family protein (TAIR:AT3G10810.1); Has 14277 Blast hits to 9531 proteins in 665 species: Archae - 26; Bacteria - 1625; Metazoa - 4928; Fungi - 2262; Plants - 3165; Viruses - 515; Other Eukaryotes - 1756 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.420.1.S1_at | 0.0602 | -1.4422 | Glyma03g37170.1 | AT5G27720.1 | embryo defective 1644 (emb1644); FUNCTIONS IN: molecular_function unknown; INVOLVED IN: embryonic development ending in seed dormancy; LOCATED IN: small nucleolar ribonucleoprotein complex, nucleus; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Like-Sm ribonucleoprotein, core (InterPro:IPR001163), Like-Sm ribonucleoprotein, eukaryotic and archaea-type, core (InterPro:IPR006649), Like-Sm ribonucleoprotein-related, core (InterPro:IPR010920); BEST Arabidopsis thaliana protein match is: small nuclear ribonucleoprotein, putative / snRNP, putative / Sm protein, putative (TAIR:AT1G20580.1); Has 1164 Blast hits to 1156 proteins in 179 species: Archae - 0; Bacteria - 6; Metazoa - 536; Fungi - 265; Plants - 172; Viruses - 5; Other Eukaryotes - 180 (source: NCBI BLink). |
| Gma.1277.3.A1_s_at | 0.08629 | -1.4397 | No Soybean Match Identified, E<10E-30, Perecent ID>95% | AT5G59370.2 | Encodes one of eight Arabidopsis actins. ACT4 belongs to the reproductive actin subclass which is predominantly expressed in developing and reproductive tissues, such as pollen, pollen tubes, ovules, and developing seeds. Expression of the ACT4/GUS fusion was restricted to young vascular tissues, tapetum, and developing and mature pollen. |

| | | | | | |
|---|---|---|---|---|---|
| Gma.2459.1.S1_at | 0.07969 | -1.4234 | Glyma06g16410.1 | AT4G30330.1 | small nuclear ribonucleoprotein E, putative / snRNP-E, putative / Sm protein E, putative; FUNCTIONS IN: molecular_function unknown; LOCATED IN: small nucleolar ribonucleoprotein complex, nucleus; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Like-Sm ribonucleoprotein, core (InterPro:IPR001163), Like-Sm ribonucleoprotein, eukaryotic and archaea-type, core (InterPro:IPR006649), Like-Sm ribonucleoprotein-related, core (InterPro:IPR010920); BEST Arabidopsis thaliana protein match is: small nuclear ribonucleoprotein E, putative / snRNP-E, putative / Sm protein E, putative (TAIR:AT2G18740.1); Has 525 Blast hits to 525 proteins in 150 species: Archae - 78; Bacteria - 0; Metazoa - 182; Fungi - 106; Plants - 70; Viruses - 0; Other Eukaryotes - 89 (source: NCBI BLink). |
| GmaAffx.79438.2.S1_at | 0.03172 | -1.4045 | No Soybean Match Identified, E<10E-30, Perecent ID>95% | ATCG00140.1 | ATPase III subunit |
| GmaAffx.89723.1.S1_x_at | 0.08729 | -1.4024 | Glyma16g01860.1 | AT5G63140.1 | ARABIDOPSIS THALIANA PURPLE ACID PHOSPHATASE 29 (ATPAP29); FUNCTIONS IN: protein serine/threonine phosphatase activity, acid phosphatase activity; LOCATED IN: endomembrane system; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Metallophosphoesterase (InterPro:IPR004843), Phosphoesterase At2g46880 (InterPro:IPR011230); BEST Arabidopsis thaliana protein match is: PAP14 (PURPLE ACID PHOSPHATASE 14); acid phosphatase/ protein serine/threonine phosphatase (TAIR:AT2G46880.1); Has 317 Blast hits to 316 proteins in 98 species: Archae - 0; Bacteria - 100; Metazoa - 0; Fungi - 132; Plants - 63; Viruses - 0; Other Eukaryotes - 22 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.1301.74.S1_s_at | 0.07811 | -1.3951 | Glyma07g13840.1 | AT3G60280.1 | Encodes blue copper-binding protein III. |
| Gma.17765.1.S1_at | 0.07969 | -1.3789 | Glyma03g26060.2 | AT2G44790.1 | Encodes a uclacyanin, a protein precursor that is closely related to precursors of stellacyanins and a blue copper protein from pea pods. |
| GmaAffx.22785.1.S1_at | 0.0481 | -1.375 | Glyma03g27510.1 | AT3G56910.1 | PLASTID-SPECIFIC 50S RIBOSOMAL PROTEIN 5 (PSRP5); FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: thylakoid, chloroplast thylakoid membrane, chloroplast, chloroplast envelope; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; Has 25 Blast hits to 25 proteins in 10 species: Archae - 0; Bacteria - 0; Metazoa - 2; Fungi - 0; Plants - 23; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| Gma.3818.1.S1_at | 0.08329 | -1.3582 | Glyma01g44630.2 | AT3G22680.1 | Encodes RNA-DIRECTED DNA METHYLATION 1 (RDM1), forming a complex with DMS3 (AT3G49250) and DRD1 (AT2G16390). This complex is termed DDR. The DDR complex is required for polymerase V transcripts and RNA-directed DNA methylation. |
| Gma.12758.1.S1_at | 0.09689 | -1.3335 | Glyma19g40620.1 | AT4G01940.1 | Encodes a protein containing the NFU domain that may be involved in iron-sulfur cluster assembly. Part of a five member gene family, more closely related to NFU2 and 3 than to NFU4 and 5. Targeted to the chloroplast. |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.72806.1.S1_at | 0.09745 | -1.3227 | Glyma11g09590.1 | AT1G31760.1 | SWIB complex BAF60b domain-containing protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 9 plant structures; EXPRESSED DURING: 6 growth stages; CONTAINS InterPro DOMAIN/s: SWIB/MDM2 (InterPro:IPR003121); BEST Arabidopsis thaliana protein match is: SWIB complex BAF60b domain-containing protein (TAIR:AT2G35605.1); Has 610 Blast hits to 576 proteins in 146 species: Archae - 0; Bacteria - 132; Metazoa - 43; Fungi - 118; Plants - 189; Viruses - 6; Other Eukaryotes - 122 (source: NCBI BLink). |
| Gma.7373.1.S1_at | 0.09223 | -1.3215 | Glyma13g20240.1 | AT2G37160.2 | transducin family protein / WD-40 repeat family protein; FUNCTIONS IN: nucleotide binding; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 20 plant structures; EXPRESSED DURING: 10 growth stages; CONTAINS InterPro DOMAIN/s: WD40 repeat-like (InterPro:IPR011046), WD40 repeat, region (InterPro:IPR017986), WD40/YVTN repeat-like (InterPro:IPR015943), WD40 repeat (InterPro:IPR001680); BEST Arabidopsis thaliana protein match is: transducin family protein / WD-40 repeat family protein (TAIR:AT3G53390.2); Has 13878 Blast hits to 7458 proteins in 380 species: Archae - 36; Bacteria - 3543; Metazoa - 4573; Fungi - 2572; Plants - 1053; Viruses - 0; Other Eukaryotes - 2101 (source: NCBI BLink). |
| GmaAffx.11697.1.S1_at | 0.09762 | -1.2466 | Glyma16g07080.1 | ATCG01280.1 | hypothetical protein |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.82533.1.S1_at | 0.05378 | -1.2132 | Glyma08g06450.1 | AT1G11390.1 | ABC1 family protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 14 plant structures; EXPRESSED DURING: M germinated pollen stage, 4 anthesis, petal differentiation and expansion stage, D bilateral stage, E expanded cotyledon stage; CONTAINS InterPro DOMAIN/s: ABC-1 (InterPro:IPR004147), Protein kinase-like (InterPro:IPR011009); BEST Arabidopsis thaliana protein match is: ABC1 family protein (TAIR:AT1G61640.1); Has 6866 Blast hits to 6818 proteins in 1101 species: Archae - 67; Bacteria - 2611; Metazoa - 374; Fungi - 362; Plants - 325; Viruses - 14; Other Eukaryotes - 3113 (source: NCBI BLink). |
| Gma.17909.1.S1_at | 0.07243 | -1.2128 | Glyma03g31030.1 | AT3G63120.1 | cyclin p1;1 (CYCP1;1); FUNCTIONS IN: cyclin-dependent protein kinase activity; INVOLVED IN: regulation of cell cycle; LOCATED IN: cellular_component unknown; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Negative regulatory factor PREG (InterPro:IPR012389), Cyclin-like (InterPro:IPR011028), Cyclin-related (InterPro:IPR013763), Cyclin, N-terminal (InterPro:IPR006671), Cyclin-related 2 (InterPro:IPR013922), Cyclin (InterPro:IPR006670); BEST Arabidopsis thaliana protein match is: CYCP4;1 (cyclin p4;1); cyclin-dependent protein kinase (TAIR:AT2G44740.1); Has 851 Blast hits to 849 proteins in 166 species: Archae - 0; Bacteria - 16; Metazoa - 221; Fungi - 291; Plants - 128; Viruses - 0; Other Eukaryotes - 195 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.91706.1.S1_at | 0.08056 | -1.1366 | Glyma08g39880.2 | AT4G12790.4 | ATP-binding family protein; FUNCTIONS IN: nucleotide binding, ATP binding; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function, ATP binding (InterPro:IPR004130); BEST Arabidopsis thaliana protein match is: QQT1 (QUATRE-QUART 1); ATP binding / nucleotide binding (TAIR:AT5G22370.2); Has 1073 Blast hits to 1063 proteins in 199 species: Archae - 79; Bacteria - 0; Metazoa - 345; Fungi - 272; Plants - 80; Viruses - 0; Other Eukaryotes - 297 (source: NCBI BLink). |

**Table 5 Candidate genes related to hormones**

| Affy ID | Q-value | Fold change | Soybean Gene | Arabidopsis homolog | Description |
|---|---|---|---|---|---|
| GmaAffx.46603.1.S1_at | 0.0273 | 5.5518 | Glyma19g27260.1 | AT5G13200.1 | GRAM domain-containing protein / ABA-responsive protein-related; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 15 plant structures; EXPRESSED DURING: 8 growth stages; CONTAINS InterPro DOMAIN/s: GRAM (InterPro:IPR004182); BEST Arabidopsis thaliana protein match is: GEM (GL2-EXPRESSION MODULATOR) (TAIR:AT2G22475.1); Has 237 Blast hits to 237 proteins in 13 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 235; Viruses - 0; Other Eukaryotes - 2 (source: NCBI BLink). |
| GmaAffx.8250.2.S1_at | 0.0268 | 4.0424 | Glyma02g16090.2 | AT5G43700.1 | Auxin inducible protein similar to transcription factors. |
| GmaAffx.2744.1.S1_at | 0.0268 | 3.5316 | Glyma04g40930.1 | AT4G00880.1 | auxin-responsive family protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: response to auxin stimulus; LOCATED IN: cellular_component unknown; EXPRESSED IN: 19 plant structures; EXPRESSED DURING: 9 growth stages; CONTAINS InterPro DOMAIN/s: Auxin responsive SAUR protein (InterPro:IPR003676); BEST Arabidopsis thaliana protein match is: auxin-responsive family protein (TAIR:AT2G46690.1); Has 736 Blast hits to 726 proteins in 21 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 735; Viruses - 0; Other Eukaryotes - 1 (source: NCBI BLink). |
| GmaAffx.87970.1.A1_at | 0.063 | 2.3181 | Glyma15g10070.1 | AT1G30040.1 | Encodes a gibberellin 2-oxidase. AtGA2OX2 expression is responsive to cytokinin and KNOX activities. |

| | | | | | |
|---|---|---|---|---|---|
| Gma.1423.1.S1_a_at | 0.0232 | 1.6451 | Glyma01g38450.1 | AT1G75080.2 | Encodes a positive regulator of the brassinosteroid (BR) signalling pathway that mediates both downstream BR responses and negative feedback regulation of BR biosynthesis. There is evidence for phosphorylation-dependent nucleocytoplasmic shuttling of BZR1. GSK3-like kinases (including BIN2), 14-3-3 proteins, and the phosphatase BSU1 seem to participate in this process. Phosphorylation also appears to affect BZR1's transcriptional activities. |
| Gma.8200.1.A1_at | 0.0815 | -7.6111 | Glyma01g38180.1 | AT3G50660.1 | Encodes a 22α hydroxylase whose reaction is a rate-limiting step in brassinosteroid biosynthetic pathway. The protein is a member of CYP90B gene family. CLM is an epi-allele with small, compressed rosette, reduced internode length, and reduced fertility, appears in selfed ddm mutant plants possibly due to loss of cytosine methylation. Transcripts accumulate in actively growing tissues, and GUS expression is negatively regulated by brassinosteroids. Localized in the endoplasmic reticulum. The in vitro expressed protein can perform the C-22 hydroxylation of a variety of C27-, C28- and C29-sterols. Cholesterol was the best substrate, followed by campesterol. Sitosterol was a poor substrate. |
| Gma.4071.1.S1_at | 0.0153 | -5.2637 | Glyma17g10050.2 | AT5G14920.1 | gibberellin-regulated family protein |
| Gma.4565.1.S1_s_at | 0.0472 | -4.7647 | Glyma19g01590.2 | AT5G14920.1 | gibberellin-regulated family protein |
| GmaAffx.28948.1.A1_at | 0.0434 | -2.2812 | Glyma13g34680.1 | AT5G20040.1 | Encodes tRNA isopentenyltransferase AtIPT9. |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.36972.2.S1_at | 0.0973 | -2.2619 | Glyma05g32990.2 | AT5G53590.1 | auxin-responsive family protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: response to auxin stimulus; LOCATED IN: cellular_component unknown; EXPRESSED IN: pollen tube; CONTAINS InterPro DOMAIN/s: Auxin responsive SAUR protein (InterPro:IPR003676); BEST Arabidopsis thaliana protein match is: auxin-responsive family protein (TAIR:AT2G46690.1); Has 708 Blast hits to 691 proteins in 38 species: Archae - 0; Bacteria - 2; Metazoa - 32; Fungi - 11; Plants - 642; Viruses - 0; Other Eukaryotes - 21 (source: NCBI BLink). |
| GmaAffx.14620.1.S1_at | 0.0798 | -2.2414 | Glyma14g40530.1 | AT1G75590.1 | auxin-responsive family protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: response to auxin stimulus; EXPRESSED IN: 7 plant structures; EXPRESSED DURING: 4 anthesis, C globular stage, petal differentiation and expansion stage; CONTAINS InterPro DOMAIN/s: Auxin responsive SAUR protein (InterPro:IPR003676); BEST Arabidopsis thaliana protein match is: auxin-responsive family protein (TAIR:AT1G19840.1); Has 730 Blast hits to 719 proteins in 19 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 729; Viruses - 0; Other Eukaryotes - 1 (source: NCBI BLink). |
| Gma.1309.1.S1_at | 0.0775 | -2.0904 | Glyma11g05710.1 | AT4G36910.1 | Has a cystathionine Beta-synthase domain. |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.36972.1.S1_at | 0.056 | -1.983 | Glyma08g00640.1 | AT5G53590.1 | auxin-responsive family protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: response to auxin stimulus; LOCATED IN: cellular_component unknown; EXPRESSED IN: pollen tube; CONTAINS InterPro DOMAIN/s: Auxin responsive SAUR protein (InterPro:IPR003676); BEST Arabidopsis thaliana protein match is: auxin-responsive family protein (TAIR:AT2G46690.1); Has 708 Blast hits to 691 proteins in 38 species: Archae - 0; Bacteria - 2; Metazoa - 32; Fungi - 11; Plants - 642; Viruses - 0; Other Eukaryotes - 21 (source: NCBI BLink). |
| Gma.4850.1.S1_at | 0.09 | -1.8694 | Glyma18g53400.1 | AT1G67080.1 | Involved in the photoprotection of PSII. aba4-1 mutant completely lacks neoxanthin,a component of the chromophore of the peripheral antenna system in PSII. Expresses neoxanthin synthase activity involved in the neoxanthin biosynthesis, an intermediary in the abscisic acid biosynthesis. |
| Gma.16437.1.S1_at | 0.0953 | -1.8312 | Glyma14g09110.1 | AT3G19270.1 | Encodes a protein with ABA 8'-hydroxylase activity, involved in ABA catabolism. Member of the CYP707A gene family. |

**Table 6 Candidate genes related to proteolysis**

| Affy ID | Q-value | Fold change | Soybean Gene | Arabidopsis homolog | Description |
|---|---|---|---|---|---|
| GmaAffx.31266.1.S1_s_at | 0.027 | 5.35364 | Glyma19g22800.1 | AT5G14420.2 | RING domain Ligase2 (RGLG2); FUNCTIONS IN: ubiquitin-protein ligase activity; INVOLVED IN: cytokinin metabolic process, auxin metabolic process; LOCATED IN: plasma membrane; EXPRESSED IN: 26 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Zinc finger, RING-type, conserved site (InterPro:IPR017907), Zinc finger, RING-type (InterPro:IPR001841), Copine (InterPro:IPR010734), von Willebrand factor, type A (InterPro:IPR002035); BEST Arabidopsis thaliana protein match is: RGLG1 (RING domain Ligase1); protein binding / zinc ion binding (TAIR:AT3G01650.1); Has 1866 Blast hits to 1803 proteins in 167 species: Archae - 0; Bacteria - 29; Metazoa - 1158; Fungi - 48; Plants - 274; Viruses - 50; Other Eukaryotes - 307 (source: NCBI BLink). |
| GmaAffx.9142.1.S1_at | 0.063 | 4.5902 | Ambiguous Hit | AT3G59940.1 | kelch repeat-containing F-box family protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Cyclin-like F-box (InterPro:IPR001810), Galactose oxidase/kelch, beta-propeller (InterPro:IPR011043), Kelch repeat type 1 (InterPro:IPR006652), Kelch related (InterPro:IPR013089), Kelch-type beta propeller (InterPro:IPR015915); BEST Arabidopsis thaliana protein match is: kelch repeat-containing F-box family protein (TAIR:AT2G44130.1); Has 2119 Blast hits to 1745 proteins in 136 species: Archae - 4; Bacteria - 165; Metazoa - 1375; Fungi - 20; Plants - 457; Viruses - 0; Other Eukaryotes - 98 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.69637.1.S1_at | 0.083 | 3.8723 | Glyma12g06860.1 | AT3G46510.1 | Encodes a protein containing a UND, a U-box, and an ARM domain. This protein has E3 ubiquitin ligase activity based on in vitro assays. |
| Gma.7923.1.A1_at | 0.078 | 3.46428 | Glyma17g09850.1 | AT1G60190.1 | armadillo/beta-catenin repeat family protein / U-box domain-containing protein; FUNCTIONS IN: ubiquitin-protein ligase activity, protein binding, binding, zinc ion binding; INVOLVED IN: protein ubiquitination; LOCATED IN: ubiquitin ligase complex; EXPRESSED IN: 6 plant structures; EXPRESSED DURING: 4 anthesis, petal differentiation and expansion stage; CONTAINS InterPro DOMAIN/s: Zinc finger, RING-type (InterPro:IPR001841), U box (InterPro:IPR003613), Zinc finger, RING/FYVE/PHD-type (InterPro:IPR013083), Armadillo-like helical (InterPro:IPR011989), Armadillo (InterPro:IPR000225), Armadillo-type fold (InterPro:IPR016024); BEST Arabidopsis thaliana protein match is: PUB18 (PLANT U-BOX 18); ubiquitin-protein ligase (TAIR:AT1G10560.1); Has 1664 Blast hits to 1609 proteins in 143 species: Archae - 2; Bacteria - 20; Metazoa - 246; Fungi - 74; Plants - 1136; Viruses - 3; Other Eukaryotes - 183 (source: NCBI BLink). |
| Gma.13446.1.A1_at | 0.048 | 2.95269 | Glyma04g41060.1 | AT2G46620.1 | AAA-type ATPase family protein; FUNCTIONS IN: nucleoside-triphosphatase activity, ATPase activity, nucleotide binding, ATP binding; LOCATED IN: endomembrane system; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: ATPase, AAA-type, core (InterPro:IPR003959), ATPase, AAA+ type, core (InterPro:IPR003593); BEST Arabidopsis thaliana protein match is: AAA-type ATPase family protein (TAIR:AT3G28580.1); Has 9837 Blast hits to 9388 proteins in 1303 species: Archae - 597; Bacteria - 2959; Metazoa - 1717; Fungi - 1281; Plants - 860; Viruses - 16; Other Eukaryotes - 2407 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.60513.1.S1_at | 0.06 | 2.89526 | Glyma07g39990.1 | AT5G59810.1 | SBT5.4; FUNCTIONS IN: identical protein binding, serine-type endopeptidase activity; INVOLVED IN: proteolysis, negative regulation of catalytic activity; EXPRESSED IN: 12 plant structures; EXPRESSED DURING: 4 anthesis, C globular stage, petal differentiation and expansion stage; CONTAINS InterPro DOMAIN/s: Protease-associated PA (InterPro:IPR003137), Proteinase inhibitor, propeptide (InterPro:IPR009020), Peptidase S8 and S53, subtilisin, kexin, sedolisin (InterPro:IPR000209), Peptidase S8, subtilisin-related (InterPro:IPR015500), Proteinase inhibitor I9, subtilisin propeptide (InterPro:IPR010259); BEST Arabidopsis thaliana protein match is: AIR3; serine-type endopeptidase (TAIR:AT2G04160.1); Has 3746 Blast hits to 3395 proteins in 608 species: Archae - 118; Bacteria - 2110; Metazoa - 79; Fungi - 147; Plants - 899; Viruses - 0; Other Eukaryotes - 393 (source: NCBI BLink). |
| Gma.5964.1.S1_at | 0.048 | 2.61819 | Ambiguous Hit | AT3G13235.1 | ubiquitin family protein; FUNCTIONS IN: aspartic-type endopeptidase activity; INVOLVED IN: response to cadmium ion; LOCATED IN: cytosol, nucleus; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Ubiquitin-associated/translation elongation factor EF1B, N-terminal, eukaryote (InterPro:IPR015940), Ubiquitin-associated/translation elongation factor EF1B, N-terminal (InterPro:IPR000449), Peptidase A2A, retrovirus, catalytic (InterPro:IPR001995), Peptidase aspartic, catalytic (InterPro:IPR009007), Ubiquitin (InterPro:IPR000626), UBA-like (InterPro:IPR009060); Has 6366 Blast hits to 3177 proteins in 509 species: Archae - 0; Bacteria - 2; Metazoa - 2716; Fungi - 819; Plants - 1543; Viruses - 81; Other Eukaryotes - 1205 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.79762.1.S1_at | 0.062 | 2.51765 | Glyma06g11990.1 | AT1G44130.1 | nucellin protein, putative; FUNCTIONS IN: aspartic-type endopeptidase activity; INVOLVED IN: proteolysis; LOCATED IN: plant-type cell wall; EXPRESSED IN: stem, embryo, pedicel, stamen; EXPRESSED DURING: 4 anthesis, C globular stage, petal differentiation and expansion stage; CONTAINS InterPro DOMAIN/s: Peptidase aspartic, catalytic (InterPro:IPR009007), Peptidase A1 (InterPro:IPR001461), Peptidase aspartic, active site (InterPro:IPR001969); BEST Arabidopsis thaliana protein match is: nucellin protein, putative (TAIR:AT1G77480.2); Has 1109 Blast hits to 1105 proteins in 75 species: Archae - 0; Bacteria - 0; Metazoa - 30; Fungi - 22; Plants - 1004; Viruses - 0; Other Eukaryotes - 53 (source: NCBI BLink). |
| GmaAffx.15783.1.A1_at | 0.027 | 2.41774 | Glyma03g31740.1 | AT4G03030.1 | kelch repeat-containing F-box family protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: chloroplast; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Cyclin-like F-box (InterPro:IPR001810), Galactose oxidase/kelch, beta-propeller (InterPro:IPR011043), Kelch repeat type 1 (InterPro:IPR006652), Kelch related (InterPro:IPR013089), Kelch-type beta propeller (InterPro:IPR015915); BEST Arabidopsis thaliana protein match is: kelch repeat-containing F-box family protein (TAIR:AT3G63220.2); Has 1405 Blast hits to 1314 proteins in 98 species: Archae - 0; Bacteria - 31; Metazoa - 962; Fungi - 4; Plants - 358; Viruses - 9; Other Eukaryotes - 41 (source: NCBI BLink). |
| GmaAffx.93148.1.S1_at | 0.055 | 2.38703 | Glyma11g35820.1 | AT3G56190.1 | Encodes one of two alpha-SNAPs (soluble NSF attachment protein) in Arabidopsis |

| | | | | | |
|---|---|---|---|---|---|
| Gma.13898.1.A1_at | 0.052 | 2.2102 | Glyma16g31980.1 | AT5G22720.1 | F-box family protein; LOCATED IN: cellular_component unknown; CONTAINS InterPro DOMAIN/s: Cyclin-like F-box (InterPro:IPR001810), FBD (InterPro:IPR013596), FBD-like (InterPro:IPR006566), Leucine-rich repeat 2 (InterPro:IPR013101); BEST Arabidopsis thaliana protein match is: F-box family protein (TAIR:AT5G22730.1); Has 1375 Blast hits to 1345 proteins in 13 species: Archae - 0; Bacteria - 0; Metazoa - 7; Fungi - 0; Plants - 1368; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| Gma.7066.1.S1_at | 0.074 | 2.15928 | Glyma08g10680.1 | AT3G26000.1 | F-box family protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; BEST Arabidopsis thaliana protein match is: F-box family protein-related (TAIR:AT3G27290.1); Has 60 Blast hits to 60 proteins in 11 species: Archae - 0; Bacteria - 0; Metazoa - 20; Fungi - 0; Plants - 39; Viruses - 0; Other Eukaryotes - 1 (source: NCBI BLink). |
| Gma.7913.1.A1_at | 0.095 | 2.02849 | Glyma11g19640.1 | AT5G22850.1 | aspartyl protease family protein; FUNCTIONS IN: aspartic-type endopeptidase activity; INVOLVED IN: proteolysis; LOCATED IN: endomembrane system; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Peptidase aspartic, catalytic (InterPro:IPR009007), Peptidase A1 (InterPro:IPR001461); BEST Arabidopsis thaliana protein match is: aspartyl protease family protein (TAIR:AT1G08210.1); Has 3460 Blast hits to 3446 proteins in 295 species: Archae - 0; Bacteria - 0; Metazoa - 1380; Fungi - 689; Plants - 1191; Viruses - 1; Other Eukaryotes - 199 (source: NCBI BLink). |
| Gma.8481.1.S1_at | 0.068 | 2.00548 | Glyma05g04230.1 | AT1G62710.1 | Encodes a vacuolar processing enzyme belonging to a novel group of cysteine proteases that is expressed specifically in seeds and is essential for the proper processing of storage proteins. |

| | | | | | |
|---|---|---|---|---|---|
| Gma.13442.1.A1_at | 0.045 | 1.98272 | Glyma12g30430.1 | AT5G22850.1 | aspartyl protease family protein; FUNCTIONS IN: aspartic-type endopeptidase activity; INVOLVED IN: proteolysis; LOCATED IN: endomembrane system; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Peptidase aspartic, catalytic (InterPro:IPR009007), Peptidase A1 (InterPro:IPR001461); BEST Arabidopsis thaliana protein match is: aspartyl protease family protein (TAIR:AT1G08210.1); Has 3460 Blast hits to 3446 proteins in 295 species: Archae - 0; Bacteria - 0; Metazoa - 1380; Fungi - 689; Plants - 1191; Viruses - 1; Other Eukaryotes - 199 (source: NCBI BLink). |
| Gma.17742.1.S1_at | 0.075 | 1.89853 | Glyma08g03630.1 | AT4G29330.1 | DERLIN-1 (DER1); FUNCTIONS IN: molecular_function unknown; INVOLVED IN: proteolysis; LOCATED IN: cellular_component unknown; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Der1-like (InterPro:IPR007599); BEST Arabidopsis thaliana protein match is: DER2.2 (DERLIN-2.2) (TAIR:AT4G04860.1); Has 666 Blast hits to 665 proteins in 170 species: Archae - 0; Bacteria - 12; Metazoa - 290; Fungi - 127; Plants - 84; Viruses - 0; Other Eukaryotes - 153 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.1753.1.S1_x_at | 0.027 | 1.88089 | Glyma08g05350.1 | AT2G32730.1 | 26S proteasome regulatory subunit, putative; FUNCTIONS IN: enzyme regulator activity, binding; INVOLVED IN: ubiquitin-dependent protein catabolic process; LOCATED IN: cytosol, proteasome regulatory particle, base subcomplex; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Armadillo-like helical (InterPro:IPR011989), Proteasome/cyclosome, regulatory subunit (InterPro:IPR002015), Armadillo-type fold (InterPro:IPR016024), 26S proteasome regulatory complex, non-ATPase subcomplex, Rpn2/Psmd1 subunit (InterPro:IPR016642); BEST Arabidopsis thaliana protein match is: 26S proteasome regulatory subunit, putative (TAIR:AT1G04810.1); Has 863 Blast hits to 795 proteins in 201 species: Archae - 10; Bacteria - 48; Metazoa - 318; Fungi - 219; Plants - 82; Viruses - 0; Other Eukaryotes - 186 (source: NCBI BLink). |
| Gma.8288.2.A1_at | 0.039 | 1.86793 | Glyma11g33880.1 | AT3G03380.1 | Encodes a putative DegP protease. |
| Gma.11023.2.S1_a_at | 0.055 | 1.85197 | Glyma01g36770.4 | AT2G17760.1 | aspartyl protease family protein; FUNCTIONS IN: aspartic-type endopeptidase activity; INVOLVED IN: proteolysis; LOCATED IN: anchored to membrane; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Peptidase aspartic, catalytic (InterPro:IPR009007), Peptidase A1 (InterPro:IPR001461), Peptidase aspartic, active site (InterPro:IPR001969); BEST Arabidopsis thaliana protein match is: aspartyl protease family protein (TAIR:AT4G35880.1); Has 1616 Blast hits to 1610 proteins in 172 species: Archae - 0; Bacteria - 0; Metazoa - 222; Fungi - 313; Plants - 988; Viruses - 0; Other Eukaryotes - 93 (source: NCBI BLink). |
| Gma.15883.1.S1_x_at | 0.095 | 1.84417 | Glyma10g01760.1 | AT3G14290.1 | Encodes 20S proteasome subunit PAE2 (PAE2). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.17881.1.S1_at | 0.039 | 1.83348 | Glyma03g39500.1 | AT3G05530.1 | Encodes RPT5a (Regulatory Particle 5a), one of the six AAA-ATPases of the proteasome regulatory particle. Essential for gametophyte development. In Arabidopsis, the RPT5 subunit is encoded by two highly homologous genes, RPT5a and RPT5b. RPT5a and RPT5b show accession-dependent functional redundancy. In Wassilewskija (Ws) accession: mutant alleles of RPT5a displayed 50% pollen lethality, indicating that RPT5a is essential for male gametophyte development. In the Columbia (Col) accession, a rpt5a mutant allele did not display such a phenotype because the RPT5b Col allele complements the rpt5a defect in the male gametophyte, whereas the RPT5b Ws allele does not. Double rpt5a rpt5b mutants in Col background showed a complete male and female gametophyte lethal phenotype. |
| Gma.1753.1.S1_a_at | 0.058 | 1.81347 | Glyma08g05350.1 | AT2G32730.1 | 26S proteasome regulatory subunit, putative; FUNCTIONS IN: enzyme regulator activity, binding; INVOLVED IN: ubiquitin-dependent protein catabolic process; LOCATED IN: cytosol, proteasome regulatory particle, base subcomplex; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Armadillo-like helical (InterPro:IPR011989), Proteasome/cyclosome, regulatory subunit (InterPro:IPR002015), Armadillo-type fold (InterPro:IPR016024), 26S proteasome regulatory complex, non-ATPase subcomplex, Rpn2/Psmd1 subunit (InterPro:IPR016642); BEST Arabidopsis thaliana protein match is: 26S proteasome regulatory subunit, putative (TAIR:AT1G04810.1); Has 863 Blast hits to 795 proteins in 201 species: Archae - 10; Bacteria - 48; Metazoa - 318; Fungi - 219; Plants - 82; Viruses - 0; Other Eukaryotes - 186 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.2885.1.S1_at | 0.045 | 1.75674 | Glyma10g31260.2 | AT2G02360.1 | Phloem protein 2-B10 (AtPP2-B10); FUNCTIONS IN: carbohydrate binding; LOCATED IN: cellular_component unknown; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Cyclin-like F-box (InterPro:IPR001810); BEST Arabidopsis thaliana protein match is: AtPP2-B8 (Phloem protein 2-B8); carbohydrate binding (TAIR:AT2G02340.1); Has 336 Blast hits to 323 proteins in 31 species: Archae - 0; Bacteria - 0; Metazoa - 4; Fungi - 0; Plants - 332; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| GmaAffx.36859.1.A1_at | 0.056 | 1.75289 | Glyma17g05490.1 | AT5G22850.1 | aspartyl protease family protein; FUNCTIONS IN: aspartic-type endopeptidase activity; INVOLVED IN: proteolysis; LOCATED IN: endomembrane system; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Peptidase aspartic, catalytic (InterPro:IPR009007), Peptidase A1 (InterPro:IPR001461); BEST Arabidopsis thaliana protein match is: aspartyl protease family protein (TAIR:AT1G08210.1); Has 3460 Blast hits to 3446 proteins in 295 species: Archae - 0; Bacteria - 0; Metazoa - 1380; Fungi - 689; Plants - 1191; Viruses - 1; Other Eukaryotes - 199 (source: NCBI BLink). |
| Gma.7865.1.S1_at | 0.08 | 1.66984 | Glyma06g12260.1 | AT3G24590.1 | Encodes a signal peptidase Plsp1 (plastidic type I signal peptidase 1). Required for thylakoid development. Functions in the maturation of the 75-kD component of the translocon at the outer envelope membrane of chloroplasts and oxygen evolving complex subunit 33 (OE33). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.1570.1.S1_at | 0.095 | 1.66335 | No Soybean Match Identified, E<10E-30, Perecent ID>95% | AT4G04860.1 | DERLIN-2.2 (DER2.2); INVOLVED IN: proteolysis; LOCATED IN: cellular_component unknown; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Der1-like (InterPro:IPR007599); BEST Arabidopsis thaliana protein match is: DER2.1 (DERLIN-2.1) (TAIR:AT4G21810.1); Has 634 Blast hits to 633 proteins in 161 species: Archae - 0; Bacteria - 0; Metazoa - 294; Fungi - 119; Plants - 84; Viruses - 0; Other Eukaryotes - 137 (source: NCBI BLink). |
| Gma.7303.1.S1_at | 0.077 | 1.63139 | Glyma06g10680.3 | AT4G08980.4 | F-box family protein (FBW2); FUNCTIONS IN: ubiquitin-protein ligase activity; INVOLVED IN: ubiquitin-dependent protein catabolic process; LOCATED IN: cellular_component unknown; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Cyclin-like F-box (InterPro:IPR001810); BEST Arabidopsis thaliana protein match is: F-box family protein (FBL20) (TAIR:AT4G05460.1); Has 817 Blast hits to 732 proteins in 67 species: Archae - 0; Bacteria - 0; Metazoa - 475; Fungi - 12; Plants - 305; Viruses - 0; Other Eukaryotes - 25 (source: NCBI BLink). |
| Gma.518.1.S1_at | 0.023 | 1.61046 | Glyma20g38970.1 | AT3G22630.1 | Encodes 20S proteasome beta subunit PBD1 (PBD1). |
| Gma.7685.1.S1_at | 0.042 | 1.60287 | Glyma06g27660.2 | AT5G57950.1 | 26S proteasome regulatory subunit, putative; FUNCTIONS IN: protein binding; INVOLVED IN: intracellular signaling cascade; LOCATED IN: proteasome regulatory particle; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: PDZ/DHR/GLGF (InterPro:IPR001478); Has 352 Blast hits to 352 proteins in 164 species: Archae - 0; Bacteria - 46; Metazoa - 112; Fungi - 85; Plants - 23; Viruses - 0; Other Eukaryotes - 86 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.11885.1.A1_a_at | 0.07 | 1.58996 | Glyma07g36390.1 | AT3G17205.1 | UBIQUITIN PROTEIN LIGASE 6 (UPL6); FUNCTIONS IN: ubiquitin-protein ligase activity; INVOLVED IN: protein modification process, protein ubiquitination; LOCATED IN: ubiquitin ligase complex, intracellular; CONTAINS InterPro DOMAIN/s: HECT (InterPro:IPR000569), IQ calmodulin-binding region (InterPro:IPR000048); BEST Arabidopsis thaliana protein match is: UPL7; ubiquitin-protein ligase (TAIR:AT3G53090.2); Has 3380 Blast hits to 3333 proteins in 174 species: Archae - 0; Bacteria - 0; Metazoa - 2123; Fungi - 505; Plants - 160; Viruses - 0; Other Eukaryotes - 592 (source: NCBI BLink). |
| Gma.18001.1.S1_at | 0.054 | 1.58377 | Glyma01g31360.3 | AT3G22110.1 | Encodes the alpha-3 subunit of 20s proteasome. |
| Gma.12102.1.S1_at | 0.042 | 1.56089 | Glyma08g45980.1 | AT3G07360.1 | Encodes a protein containing a U-box and an ARM domain. This protein has E3 ubiquitin ligase activity based on in vitro assays. |
| Gma.4091.2.S1_at | 0.068 | 1.55231 | Glyma13g27930.1 | AT3G06140.1 | zinc finger (C3HC4-type RING finger) family protein; FUNCTIONS IN: protein binding, zinc ion binding; EXPRESSED IN: 17 plant structures; EXPRESSED DURING: 8 growth stages; CONTAINS InterPro DOMAIN/s: Zinc finger, RING-type (InterPro:IPR001841), Zinc finger, RING/FYVE/PHD-type (InterPro:IPR013083); BEST Arabidopsis thaliana protein match is: zinc finger (C3HC4-type RING finger) family protein (TAIR:AT5G19080.1); Has 6615 Blast hits to 4433 proteins in 311 species: Archae - 2; Bacteria - 83; Metazoa - 2278; Fungi - 466; Plants - 2616; Viruses - 263; Other Eukaryotes - 907 (source: NCBI BLink). |
| GmaAffx.21553.1.S1_at | 0.054 | 1.55052 | Glyma08g40300.1 | AT1G70170.1 | mutant has Late flowering; Early senescence; Matrix Metalloproteinase |
| Gma.10691.1.S1_at | 0.058 | 1.54548 | Glyma03g41630.2 | AT4G27960.2 | ubiquitin conjugating enzyme |
| Gma.10883.2.S1_s_at | 0.096 | 1.47941 | Glyma13g17820.1 | AT5G20620.1 | encodes a ubiquitin polyprotein. |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.74168.1.S1_at | 0.088 | 1.47672 | Glyma14g40620.1 | AT4G24160.1 | Encodes a soluble lysophosphatidic acid acyltransferase with additional triacylglycerol lipase and phosphatidylcholine hydrolyzing enzymatic activities. Plays a pivotal role in maintaining the lipid homeostasis by regulating both phospholipid and neutral lipid levels. |
| GmaAffx.54705.1.S1_at | 0.078 | 1.46388 | Glyma06g44470.1 | AT5G44280.1 | Encodes a nuclear localized protein with similarity to animal polycomb repressive core complex1 (PRC1) core component RING.Appears to function redundantly with ATRING1b, a close paralog. Both interact physically with CLF and LHP1 and appear to function together to repress class I KNOX gene expression. |
| Gma.6872.1.S1_at | 0.034 | 1.45512 | Glyma12g36650.2 | AT3G63530.2 | Encodes a novel E3 ubiquitin ligase that acts as a central negative regulator of Arabidopsis floral organ size. Organ size is limited by restricting the period of proliferative growth, rather than the rate of growth and appears to act in a novel pathway, independent of ANT and JAG. Both mRNA and protein are expressed in all actively growing regions of the plant and the vasculature and the protein is rapidly turned over by proteasome-mediated degradation. |
| Gma.5235.1.S1_at | 0.06 | 1.45318 | Glyma12g03250.1 | AT5G48655.3 | zinc finger (C3HC4-type RING finger) family protein; FUNCTIONS IN: protein binding, zinc ion binding; INVOLVED IN: response to chitin; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Zinc finger, RING-type, conserved site (InterPro:IPR017907), Zinc finger, RING-type (InterPro:IPR001841), Zinc finger, RING/FYVE/PHD-type (InterPro:IPR013083); BEST Arabidopsis thaliana protein match is: zinc finger (C3HC4-type RING finger) family protein (TAIR:AT3G07200.2); Has 3740 Blast hits to 3729 proteins in 236 species: Archae - 0; Bacteria - 0; Metazoa - 2631; Fungi - 439; Plants - 221; Viruses - 18; Other Eukaryotes - 431 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.286.1.S1_at | 0.071 | 1.412 | No Soybean Match Identified, E<10E-30, Perecent ID>95% | AT1G62290.2 | aspartyl protease family protein; FUNCTIONS IN: aspartic-type endopeptidase activity; INVOLVED IN: proteolysis, lipid metabolic process; LOCATED IN: vacuole; EXPRESSED IN: 10 plant structures; EXPRESSED DURING: LP.06 six leaves visible, LP.04 four leaves visible, seedling growth, petal differentiation and expansion stage, LP.08 eight leaves visible; CONTAINS InterPro DOMAIN/s: Saposin-like (InterPro:IPR011001), Peptidase aspartic, catalytic (InterPro:IPR009007), Saposin-like type B, 1 (InterPro:IPR007856), Saposin-like type B, 2 (InterPro:IPR008138), Saposin B (InterPro:IPR008139), Peptidase A1 (InterPro:IPR001461), Peptidase aspartic, active site (InterPro:IPR001969); BEST Arabidopsis thaliana protein match is: aspartyl protease family protein (TAIR:AT1G11910.1); Has 5754 Blast hits to 3941 proteins in 342 species: Archae - 0; Bacteria - 4; Metazoa - 3397; Fungi - 1194; Plants - 356; Viruses - 0; Other Eukaryotes - 803 (source: NCBI BLink). |
| Gma.17082.1.S1_at | 0.048 | 1.36975 | Glyma11g34460.2 | AT2G39760.1 | BPM3; FUNCTIONS IN: protein binding; INVOLVED IN: biological_process unknown; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: TRAF-like (InterPro:IPR008974), MATH (InterPro:IPR002083), BTB/POZ fold (InterPro:IPR011333), BTB/POZ (InterPro:IPR013069), Kelch related (InterPro:IPR013089), BTB/POZ-like (InterPro:IPR000210), TRAF-type (InterPro:IPR013322); BEST Arabidopsis thaliana protein match is: ATBPM1 (BTB-POZ and MATH domain 1); protein binding (TAIR:AT5G19000.1); Has 5596 Blast hits to 5441 proteins in 138 species: Archae - 0; Bacteria - 0; Metazoa - 4431; Fungi - 78; Plants - 862; Viruses - 29; Other Eukaryotes - 196 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.86415.1.S1_at | 0.06 | 1.3649 | Glyma09g04090.1 | AT3G17000.1 | ubiquitin-conjugating enzyme 32 (UBC32); FUNCTIONS IN: ubiquitin-protein ligase activity; INVOLVED IN: ubiquitin-dependent protein catabolic process; LOCATED IN: cellular_component unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Ubiquitin-conjugating enzyme/RWD-like (InterPro:IPR016135), Ubiquitin-conjugating enzyme, E2 (InterPro:IPR000608); BEST Arabidopsis thaliana protein match is: UBC34 (ubiquitin-conjugating enzyme 34); ubiquitin-protein ligase (TAIR:AT1G17280.2); Has 5810 Blast hits to 5810 proteins in 294 species: Archae - 0; Bacteria - 0; Metazoa - 2855; Fungi - 1053; Plants - 896; Viruses - 16; Other Eukaryotes - 990 (source: NCBI BLink). |
| GmaAffx.57400.1.A1_at | 0.063 | -6.3978 | No Soybean Match Identified, E<10E-30, Perecent ID>95% | AT1G28110.2 | SERINE CARBOXYPEPTIDASE-LIKE 45 PRECURSOR (SCPL45); FUNCTIONS IN: serine-type carboxypeptidase activity; INVOLVED IN: proteolysis; LOCATED IN: plant-type cell wall; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Peptidase S10, serine carboxypeptidase (InterPro:IPR001563), Peptidase S10, serine carboxypeptidase, active site (InterPro:IPR018202); BEST Arabidopsis thaliana protein match is: scpl46 (serine carboxypeptidase-like 46); serine-type carboxypeptidase (TAIR:AT2G33530.1); Has 2581 Blast hits to 2531 proteins in 335 species: Archae - 0; Bacteria - 249; Metazoa - 573; Fungi - 560; Plants - 874; Viruses - 0; Other Eukaryotes - 325 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.55671.2.S1_at | 0.095 | -3.9231 | Glyma16g24270.1 | AT4G36195.1 | serine carboxypeptidase S28 family protein; FUNCTIONS IN: serine-type peptidase activity; INVOLVED IN: proteolysis; LOCATED IN: plasma membrane, vacuole, plant-type cell wall; EXPRESSED IN: stem, guard cell, cultured cell; CONTAINS InterPro DOMAIN/s: Peptidase S28 (InterPro:IPR008758); BEST Arabidopsis thaliana protein match is: serine-type peptidase (TAIR:AT4G36190.1); Has 912 Blast hits to 876 proteins in 118 species: Archae - 0; Bacteria - 16; Metazoa - 518; Fungi - 136; Plants - 101; Viruses - 0; Other Eukaryotes - 141 (source: NCBI BLink). |
| GmaAffx.20786.1.A1_at | 0.045 | -3.442 | Glyma17g04120.2 | AT3G17180.1 | serine carboxypeptidase-like 33 (scpl33); FUNCTIONS IN: serine-type carboxypeptidase activity; INVOLVED IN: proteolysis; LOCATED IN: endomembrane system; EXPRESSED IN: shoot apex, flower, root, stamen; EXPRESSED DURING: 4 anthesis, petal differentiation and expansion stage; CONTAINS InterPro DOMAIN/s: Peptidase S10, serine carboxypeptidase (InterPro:IPR001563), Peptidase S10, serine carboxypeptidase, active site (InterPro:IPR018202); BEST Arabidopsis thaliana protein match is: SCPL34; serine-type carboxypeptidase (TAIR:AT5G23210.1); Has 2555 Blast hits to 2510 proteins in 313 species: Archae - 0; Bacteria - 191; Metazoa - 568; Fungi - 566; Plants - 899; Viruses - 0; Other Eukaryotes - 331 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.20786.2.S1_at | 0.05 | -2.661 | Glyma17g04120.2 | AT3G17180.1 | serine carboxypeptidase-like 33 (scpl33); FUNCTIONS IN: serine-type carboxypeptidase activity; INVOLVED IN: proteolysis; LOCATED IN: endomembrane system; EXPRESSED IN: shoot apex, flower, root, stamen; EXPRESSED DURING: 4 anthesis, petal differentiation and expansion stage; CONTAINS InterPro DOMAIN/s: Peptidase S10, serine carboxypeptidase (InterPro:IPR001563), Peptidase S10, serine carboxypeptidase, active site (InterPro:IPR018202); BEST Arabidopsis thaliana protein match is: SCPL34; serine-type carboxypeptidase (TAIR:AT5G23210.1); Has 2555 Blast hits to 2510 proteins in 313 species: Archae - 0; Bacteria - 191; Metazoa - 568; Fungi - 566; Plants - 899; Viruses - 0; Other Eukaryotes - 331 (source: NCBI BLink). |
| Gma.879.1.A1_a_at | 0.009 | -2.5896 | Glyma01g01000.2 | AT5G19740.1 | peptidase M28 family protein; FUNCTIONS IN: dipeptidase activity; INVOLVED IN: proteolysis; LOCATED IN: vacuole; CONTAINS InterPro DOMAIN/s: Protease-associated PA (InterPro:IPR003137), Transferrin receptor-like, dimerisation (InterPro:IPR007365), Peptidase M28 (InterPro:IPR007484); BEST Arabidopsis thaliana protein match is: AMP1 (ALTERED MERISTEM PROGRAM 1); carboxypeptidase/ dipeptidase (TAIR:AT3G54720.1); Has 2509 Blast hits to 2472 proteins in 344 species: Archae - 16; Bacteria - 796; Metazoa - 572; Fungi - 328; Plants - 72; Viruses - 0; Other Eukaryotes - 725 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.1620.1.S1_at | 0.064 | -2.3747 | Glyma03g28210.1 | AT2G32560.1 | F-box family protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: endomembrane system; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Cyclin-like F-box (InterPro:IPR001810); BEST Arabidopsis thaliana protein match is: F-box family protein (TAIR:AT2G26850.1); Has 52 Blast hits to 52 proteins in 12 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 51; Viruses - 0; Other Eukaryotes - 1 (source: NCBI BLink). |
| Gma.4867.1.S1_at | 0.092 | -2.0916 | Glyma03g28080.3 | AT1G28110.2 | SERINE CARBOXYPEPTIDASE-LIKE 45 PRECURSOR (SCPL45); FUNCTIONS IN: serine-type carboxypeptidase activity; INVOLVED IN: proteolysis; LOCATED IN: plant-type cell wall; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Peptidase S10, serine carboxypeptidase (InterPro:IPR001563), Peptidase S10, serine carboxypeptidase, active site (InterPro:IPR018202); BEST Arabidopsis thaliana protein match is: scpl46 (serine carboxypeptidase-like 46); serine-type carboxypeptidase (TAIR:AT2G33530.1); Has 2581 Blast hits to 2531 proteins in 335 species: Archae - 0; Bacteria - 249; Metazoa - 573; Fungi - 560; Plants - 874; Viruses - 0; Other Eukaryotes - 325 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.16526.1.S1_at | 0.066 | -1.9568 | Glyma13g03850.1 | AT3G12203.1 | serine carboxypeptidase-like 17 (scpl17); FUNCTIONS IN: serine-type carboxypeptidase activity; INVOLVED IN: proteolysis; LOCATED IN: endomembrane system; CONTAINS InterPro DOMAIN/s: Peptidase S10, serine carboxypeptidase (InterPro:IPR001563), Peptidase S10, serine carboxypeptidase, active site (InterPro:IPR018202); BEST Arabidopsis thaliana protein match is: scpl16 (serine carboxypeptidase-like 16); serine-type carboxypeptidase (TAIR:AT3G12220.1); Has 2528 Blast hits to 2469 proteins in 280 species: Archae - 0; Bacteria - 129; Metazoa - 563; Fungi - 548; Plants - 983; Viruses - 0; Other Eukaryotes - 305 (source: NCBI BLink). |
| GmaAffx.56692.1.S1_at | 0.063 | -1.7793 | Glyma07g07610.1 | AT1G01650.1 | aspartic-type endopeptidase/ peptidase; FUNCTIONS IN: peptidase activity, aspartic-type endopeptidase activity; INVOLVED IN: proteolysis; LOCATED IN: endomembrane system, integral to membrane; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Protease-associated PA (InterPro:IPR003137), Peptidase A22, presenilin signal peptide (InterPro:IPR006639), Peptidase A22B, signal peptide peptidase (InterPro:IPR007369); BEST Arabidopsis thaliana protein match is: protease-associated (PA) domain-containing protein (TAIR:AT1G63690.1); Has 1108 Blast hits to 1084 proteins in 196 species: Archae - 0; Bacteria - 119; Metazoa - 534; Fungi - 104; Plants - 174; Viruses - 0; Other Eukaryotes - 177 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.7221.1.S1_at | 0.1 | -1.6834 | Glyma13g33010.1 | AT1G09730.1 | Ulp1 protease family protein; FUNCTIONS IN: cysteine-type peptidase activity; INVOLVED IN: proteolysis; LOCATED IN: cellular_component unknown; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Peptidase C48, SUMO/Sentrin/Ubl1 (InterPro:IPR003653); BEST Arabidopsis thaliana protein match is: Ulp1 protease family protein (TAIR:AT4G33620.1); Has 1334 Blast hits to 1209 proteins in 162 species: Archae - 2; Bacteria - 12; Metazoa - 733; Fungi - 219; Plants - 125; Viruses - 7; Other Eukaryotes - 236 (source: NCBI BLink). |
| GmaAffx.8646.1.S1_at | 0.076 | -1.6495 | Glyma20g32790.2 | AT5G04710.1 | aspartyl aminopeptidase, putative; FUNCTIONS IN: aminopeptidase activity, zinc ion binding; INVOLVED IN: proteolysis; LOCATED IN: chloroplast, vacuole; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Peptidase M18, aminopeptidase I (InterPro:IPR001948); BEST Arabidopsis thaliana protein match is: aspartyl aminopeptidase, putative (TAIR:AT5G60160.1); Has 1276 Blast hits to 1273 proteins in 389 species: Archae - 1; Bacteria - 687; Metazoa - 123; Fungi - 173; Plants - 40; Viruses - 0; Other Eukaryotes - 252 (source: NCBI BLink). |
| Gma.5246.2.S1_at | 0.055 | -1.5497 | Glyma10g35630.1 | AT1G50490.1 | Encodes one of two ubiquitin-conjugating enzymes belonging to the E2-C gene family (the other being UBC19). Transcript is always found in diving cells, but also in other non-dividing cells. |

| | | | | | |
|---|---|---|---|---|---|
| Gma.3073.1.S1_at | 0.074 | -1.4923 | Glyma15g06330.1 | AT1G09730.1 | Ulp1 protease family protein; FUNCTIONS IN: cysteine-type peptidase activity; INVOLVED IN: proteolysis; LOCATED IN: cellular_component unknown; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Peptidase C48, SUMO/Sentrin/Ubl1 (InterPro:IPR003653); BEST Arabidopsis thaliana protein match is: Ulp1 protease family protein (TAIR:AT4G33620.1); Has 1334 Blast hits to 1209 proteins in 162 species: Archae - 2; Bacteria - 12; Metazoa - 733; Fungi - 219; Plants - 125; Viruses - 7; Other Eukaryotes - 236 (source: NCBI BLink). |
| Gma.3195.1.S1_at | 0.078 | -1.4367 | Glyma03g37810.1 | AT2G47390.1 | serine-type endopeptidase/ serine-type peptidase; FUNCTIONS IN: serine-type peptidase activity, serine-type endopeptidase activity; INVOLVED IN: proteolysis; LOCATED IN: chloroplast stroma, chloroplast; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Peptidase S9, prolyl oligopeptidase active site region (InterPro:IPR001375), Peptidase S9A, oligopeptidase, N-terminal beta-propeller (InterPro:IPR004106); Has 3223 Blast hits to 3222 proteins in 462 species: Archae - 101; Bacteria - 1623; Metazoa - 419; Fungi - 20; Plants - 75; Viruses - 0; Other Eukaryotes - 985 (source: NCBI BLink). |
| GmaAffx.91367.1.S1_s_at | 0.063 | -1.2848 | No Soybean Match Identified, E<10E-30, Perecent ID>95% | AT5G53300.2 | Encodes a ubiquitin conjugating enzyme. |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.79501.1.A1_at | 0.087 | -1.2633 | Glyma08g42730.1 | AT5G36210.1 | serine-type peptidase; FUNCTIONS IN: serine-type peptidase activity; INVOLVED IN: proteolysis; LOCATED IN: chloroplast; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Peptidase S9, prolyl oligopeptidase active site region (InterPro:IPR001375); BEST Arabidopsis thaliana protein match is: acylaminoacyl-peptidase-related (TAIR:AT4G14570.1); Has 5320 Blast hits to 5314 proteins in 668 species: Archae - 124; Bacteria - 2501; Metazoa - 652; Fungi - 161; Plants - 110; Viruses - 0; Other Eukaryotes - 1772 (source: NCBI BLink). |
| GmaAffx.60900.1.S1_at | 0.052 | -1.1868 | Glyma13g16580.1 | AT2G19430.1 | This gene is predicted to encode a protein with a DWD motif. It can bind to DDB1a in Y2H assays, and DDB1b in co-IP assays, and may be involved in the formation of a CUL4-based E3 ubiquitin ligase. |

**Table 7 Candidate genes related to signaling**

| Affy ID | Q-value | Fold change | Soybean Gene | Arabidopsis homolog | Description |
|---|---|---|---|---|---|
| Gma.15972.1.A1_at | 0.0232 | 27.6472 | Glyma04g17710.1 | AT5G42380.1 | CALMODULIN LIKE 37 (CML37); FUNCTIONS IN: calcium ion binding; INVOLVED IN: response to ozone; LOCATED IN: chloroplast; EXPRESSED IN: 12 plant structures; EXPRESSED DURING: 8 growth stages; CONTAINS InterPro DOMAIN/s: EF-HAND 1 (InterPro:IPR018247), EF-HAND 2 (InterPro:IPR018249), EF-Hand type (InterPro:IPR011992), Calcium-binding EF-hand (InterPro:IPR002048), EF hand (InterPro:IPR018248); BEST Arabidopsis thaliana protein match is: calmodulin-related protein, putative (TAIR:AT1G76640.1); Has 13184 Blast hits to 10239 proteins in 1127 species: Archae - 0; Bacteria - 39; Metazoa - 6469; Fungi - 2027; Plants - 2539; Viruses - 2; Other Eukaryotes - 2108 (source: NCBI BLink). |
| GmaAffx.25470.1.S1_at | 0.0872 | 13.8202 | Glyma17g37750.1 | AT1G75750.1 | GA-responsive GAST1 protein homolog regulated by BR and GA antagonistically. Possibly involved in cell elongation based on expression data |
| GmaAffx.68498.1.S1_at | 0.0268 | 13.0839 | Glyma12g11210.1 | AT4G27280.1 | calcium-binding EF hand family protein; FUNCTIONS IN: calcium ion binding; INVOLVED IN: biological_process unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: EF-HAND 1 (InterPro:IPR018247), EF-HAND 2 (InterPro:IPR018249), EF-Hand type (InterPro:IPR011992); BEST Arabidopsis thaliana protein match is: PBP1 (PINOID-BINDING PROTEIN 1); calcium ion binding / protein binding (TAIR:AT5G54490.1); Has 486 Blast hits to 485 proteins in 125 species: Archae - 0; Bacteria - 0; Metazoa - 216; Fungi - 30; Plants - 142; Viruses - 0; Other Eukaryotes - 98 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.13672.1.A1_at | 0.0595 | 10.4317 | Glyma18g03090.1 | AT3G08510.2 | Phosphoinositide-specific phospholipase C (PI-PLC), catalyzes hydrolysis of phosphatidylinositol 4,5-bisphosphate into inositol 1,4,5-trisphosphate and diacylglycerol. |
| Gma.15997.1.S1_at | 0.0232 | 9.89091 | Glyma15g06130.1 | AT5G64310.1 | Encodes arabinogalactan-protein (AGP1). |
| Gma.221.1.S1_at | 0.0595 | 6.10947 | Glyma11g35310.1 | AT3G08510.2 | Phosphoinositide-specific phospholipase C (PI-PLC), catalyzes hydrolysis of phosphatidylinositol 4,5-bisphosphate into inositol 1,4,5-trisphosphate and diacylglycerol. |
| Gma.3440.2.S1_a_at | 0.069 | 5.84875 | Glyma15g38010.2 | AT5G12380.1 | annexin, putative; FUNCTIONS IN: calcium-dependent phospholipid binding, calcium ion binding |
| GmaAffx.70425.1.S1_at | 0.0527 | 5.12627 | Glyma04g27810.1 | AT2G38800.1 | calmodulin-binding protein-related; FUNCTIONS IN: calmodulin binding |
| GmaAffx.11858.1.A1_s_at | 0.0777 | 4.89498 | Glyma12g03450.1 | AT4G38690.1 | 1-phosphatidylinositol phosphodiesterase-related; FUNCTIONS IN: phospholipase C activity, phosphoric diester hydrolase activity; INVOLVED IN: intracellular signaling cascade, N-terminal protein myristoylation, lipid metabolic process; LOCATED IN: cellular_component unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 12 growth stages; CONTAINS InterPro DOMAIN/s: Phospholipase C, phosphatidylinositol-specific , X region (InterPro:IPR000909), PLC-like phosphodiesterase, TIM beta/alpha-barrel domain (InterPro:IPR017946); BEST Arabidopsis thaliana protein match is: 1-phosphatidylinositol phosphodiesterase-related (TAIR:AT4G34920.1); Has 491 Blast hits to 491 proteins in 117 species: Archae - 0; Bacteria - 349; Metazoa - 45; Fungi - 48; Plants - 31; Viruses - 0; Other Eukaryotes - 18 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.3111.3.S1_at | 0.0772 | 3.78656 | Glyma12g06760.2 | AT2G28930.3 | APK1B; FUNCTIONS IN: protein serine/threonine kinase activity, protein kinase activity, kinase activity, ATP binding; INVOLVED IN: protein amino acid phosphorylation; LOCATED IN: plasma membrane; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Protein kinase, ATP binding site (InterPro:IPR017441), Protein kinase, core (InterPro:IPR000719), Serine/threonine protein kinase-related (InterPro:IPR017442), Protein kinase-like (InterPro:IPR011009), Serine/threonine protein kinase, active site (InterPro:IPR008271); BEST Arabidopsis thaliana protein match is: APK1A; kinase/ protein serine/threonine kinase (TAIR:AT1G07570.2); Has 83141 Blast hits to 82181 proteins in 3045 species: Archae - 42; Bacteria - 7282; Metazoa - 37065; Fungi - 6189; Plants - 18593; Viruses - 340; Other Eukaryotes - 13630 (source: NCBI BLink). |
| GmaAffx.85644.1.A1_at | 0.0841 | 3.50749 | Glyma03g27290.2 | AT3G56860.2 | encodes a nuclear protein that binds to RNA with a specificity for oligouridylates in vitro. Along with UBP1 and UBA1a, it may act as a component of a complex recognizing U-rich sequences in plant 3'-UTRs and contributing to the stabilization of mRNAs in the nucleus. |
| Gma.3223.1.S1_at | 0.0536 | 3.45703 | Glyma10g32190.1 | AT3G22930.1 | calmodulin, putative; FUNCTIONS IN: calcium ion binding |
| Gma.8586.1.S1_at | 0.0338 | 3.38531 | Glyma17g17330.2 | AT5G67600.1 | unknown protein; LOCATED IN: plasma membrane; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; Has 2474 Blast hits to 1458 proteins in 244 species: Archae - 0; Bacteria - 299; Metazoa - 816; Fungi - 328; Plants - 498; Viruses - 1; Other Eukaryotes - 532 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.56832.1.A1_at | 0.0472 | 3.04664 | Glyma17g11540.1 | AT4G32060.1 | calcium-binding EF hand family protein; FUNCTIONS IN: calcium ion binding; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: EF-HAND 2 (InterPro:IPR018249), EF-Hand type (InterPro:IPR011992), Calcium-binding EF-hand (InterPro:IPR002048), EF hand (InterPro:IPR018248); Has 902 Blast hits to 878 proteins in 119 species: Archae - 0; Bacteria - 0; Metazoa - 729; Fungi - 55; Plants - 46; Viruses - 0; Other Eukaryotes - 72 (source: NCBI BLink). |
| Gma.7064.2.S1_at | 0.0447 | 2.85506 | Glyma07g37080.1 | AT4G02080.1 | A member of ARF-like GTPase family. A thaliana has 21 members, in two subfamilies, ARF and ARF-like (ARL) GTPases. |
| Gma.13830.1.A1_at | 0.0685 | 2.63964 | Glyma08g27710.1 | AT5G28680.1 | Receptor-like kinase required for maintenance of pollen tube growth. Display polar localization at the plasma membrane of the pollen tube tip. |
| Gma.4590.1.S1_at | 0.0538 | 2.63554 | Glyma10g40550.1 | AT4G33920.1 | protein phosphatase 2C family protein / PP2C family protein; FUNCTIONS IN: protein serine/threonine phosphatase activity |
| Gma.17843.1.S1_at | 0.0455 | 2.61659 | Glyma04g02660.1 | AT1G75750.1 | GA-responsive GAST1 protein homolog regulated by BR and GA antagonistically. Possibly involved in cell elongation based on expression data |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.78772.1.S1_at | 0.0955 | 2.52469 | Glyma15g09490.2 | AT4G18950.1 | ankyrin protein kinase, putative; FUNCTIONS IN: protein serine/threonine/tyrosine kinase activity, kinase activity; INVOLVED IN: regulation of signal transduction, protein amino acid phosphorylation; LOCATED IN: nucleus, cytoplasm; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Protein kinase, core (InterPro:IPR000719), Integrin-linked protein kinase (InterPro:IPR016253), Tyrosine protein kinase (InterPro:IPR001245), Serine/threonine protein kinase-related (InterPro:IPR017442), Protein kinase-like (InterPro:IPR011009), Ankyrin (InterPro:IPR002110); BEST Arabidopsis thaliana protein match is: ankyrin protein kinase, putative (TAIR:AT3G58760.1); Has 116131 Blast hits to 99643 proteins in 3446 species: Archae - 96; Bacteria - 9447; Metazoa - 52988; Fungi - 8496; Plants - 19258; Viruses - 634; Other Eukaryotes - 25212 (source: NCBI BLink). |
| GmaAffx.45530.1.S1_at | 0.0391 | 2.50746 | Glyma11g32050.1 | AT1G70530.1 | Encodes a cysteine-rich receptor-like protein kinase. |
| Gma.7064.1.S1_at | 0.0268 | 2.44285 | Glyma07g37080.2 | AT3G62560.1 | GTP-binding protein, putative; FUNCTIONS IN: GTP binding; INVOLVED IN: intracellular protein transport; LOCATED IN: endomembrane system, intracellular; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Small GTP-binding protein (InterPro:IPR005225), GTP-binding protein SAR1 (InterPro:IPR006687), ARF/SAR superfamily (InterPro:IPR006689); BEST Arabidopsis thaliana protein match is: ATSAR2 (ARABIDOPSIS THALIANA SECRETION-ASSOCIATED RAS SUPER FAMILY 2); GTP binding (TAIR:AT4G02080.1); Has 5231 Blast hits to 5228 proteins in 291 species: Archae - 2; Bacteria - 39; Metazoa - 2741; Fungi - 852; Plants - 681; Viruses - 0; Other Eukaryotes - 916 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.4531.3.S1_at | 0.0677 | 2.3187 | Glyma04g10880.1 | AT4G08950.1 | EXORDIUM (EXO); FUNCTIONS IN: molecular_function unknown; INVOLVED IN: response to brassinosteroid stimulus; LOCATED IN: cell wall, plant-type cell wall; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Phosphate-induced protein 1 conserved region (InterPro:IPR006766); BEST Arabidopsis thaliana protein match is: PHI-1 (PHOSPHATE-INDUCED 1) (TAIR:AT1G35140.1); Has 234 Blast hits to 233 proteins in 17 species: Archae - 0; Bacteria - 2; Metazoa - 0; Fungi - 0; Plants - 232; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| Gma.3993.1.S1_at | 0.0273 | 2.28849 | Glyma11g09620.2 | AT4G19200.1 | proline-rich family protein; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: XYPPX repeat (InterPro:IPR006031); Has 21331 Blast hits to 10933 proteins in 746 species: Archae - 2; Bacteria - 4207; Metazoa - 10453; Fungi - 2106; Plants - 2404; Viruses - 136; Other Eukaryotes - 2023 (source: NCBI BLink). |
| GmaAffx.18096.1.S1_at | 0.0605 | 2.19456 | Glyma03g37760.1 | AT2G47430.1 | Encodes a putative plasma membrane-bound hybrid histidine kinase and cytokinin sensor that is expressed within the female gametophyte. |
| GmaAffx.21079.1.A1_at | 0.0232 | 2.15449 | Glyma03g32400.1 | AT5G43980.1 | Encodes a plasmodesmal protein that affects the intercellular movement of molecules through the plasmodesmata. The cytoplasmic C-terminal portion of the protein is connected to the apoplastic N-terminal portion of the protein by a single transmembrane domain (TMD). It is transported to the plasmodesmata through the secretory pathway. PDLP1 has two DUF26 domains and a signal peptide, but the proper localization of the protein appears to depend on the TMD. |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.5689.1.S1_at | 0.0894 | 2.103 | Glyma17g16000.2 | AT5G47070.1 | protein kinase, putative; FUNCTIONS IN: protein serine/threonine kinase activity, protein kinase activity, kinase activity, ATP binding; INVOLVED IN: protein amino acid phosphorylation; LOCATED IN: plasma membrane; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Protein kinase, ATP binding site (InterPro:IPR017441), Protein kinase, core (InterPro:IPR000719), Serine/threonine protein kinase-related (InterPro:IPR017442), Protein kinase-like (InterPro:IPR011009), Serine/threonine protein kinase, active site (InterPro:IPR008271); BEST Arabidopsis thaliana protein match is: protein kinase, putative (TAIR:AT4G17660.1); Has 85578 Blast hits to 84517 proteins in 3070 species: Archae - 57; Bacteria - 8047; Metazoa - 37348; Fungi - 6649; Plants - 18542; Viruses - 344; Other Eukaryotes - 14591 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.8459.1.S1_at | 0.0982 | 2.09909 | Glyma07g36200.2 | AT3G17410.1 | serine/threonine protein kinase, putative; FUNCTIONS IN: protein serine/threonine kinase activity, protein tyrosine kinase activity, protein kinase activity, kinase activity, ATP binding; INVOLVED IN: protein amino acid phosphorylation, N-terminal protein myristoylation; LOCATED IN: plasma membrane; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Protein kinase, ATP binding site (InterPro:IPR017441), Protein kinase, core (InterPro:IPR000719), Tyrosine protein kinase (InterPro:IPR001245), Protein kinase-like (InterPro:IPR011009), Serine/threonine protein kinase, active site (InterPro:IPR008271); BEST Arabidopsis thaliana protein match is: serine/threonine protein kinase, putative (TAIR:AT1G48210.2); Has 80836 Blast hits to 79972 proteins in 3240 species: Archae - 50; Bacteria - 7423; Metazoa - 35632; Fungi - 5924; Plants - 18246; Viruses - 310; Other Eukaryotes - 13251 (source: NCBI BLink). |
| Gma.7768.1.A1_s_at | 0.0566 | 2.08346 | Glyma18g41820.1 | AT3G54020.1 | Inositol phosphorylceramide synthase |

| | | | | | |
|---|---|---|---|---|---|
| Gma.5550.1.S1_at | 0.0587 | 2.08001 | Glyma16g28640.1 | AT5G42140.1 | zinc finger protein, putative / regulator of chromosome condensation (RCC1) family protein; FUNCTIONS IN: chromatin binding, phosphoinositide binding, zinc ion binding, Ran GTPase binding; INVOLVED IN: signal transduction; EXPRESSED IN: 18 plant structures; EXPRESSED DURING: 9 growth stages; CONTAINS InterPro DOMAIN/s: Regulator of chromosome condensation, RCC1 (InterPro:IPR000408), Disease resistance/zinc finger/chromosome condensation-like region (InterPro:IPR013591), Zinc finger, FYVE-type (InterPro:IPR000306), Regulator of chromosome condensation/beta-lactamase-inhibitor protein II (InterPro:IPR009091), Zinc finger, FYVE-related (InterPro:IPR017455), Pleckstrin homology-type (InterPro:IPR011993), Zinc finger, FYVE/PHD-type (InterPro:IPR011011); BEST Arabidopsis thaliana protein match is: PRAF1; Ran GTPase binding / chromatin binding / zinc ion binding (TAIR:AT1G76950.1); Has 16862 Blast hits to 6456 proteins in 315 species: Archae - 38; Bacteria - 1318; Metazoa - 8172; Fungi - 719; Plants - 1471; Viruses - 3; Other Eukaryotes - 5141 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.7064.1.S1_a_at | 0.0472 | 2.0391 | Glyma07g37080.2 | AT3G62560.1 | GTP-binding protein, putative; FUNCTIONS IN: GTP binding; INVOLVED IN: intracellular protein transport; LOCATED IN: endomembrane system, intracellular; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Small GTP-binding protein (InterPro:IPR005225), GTP-binding protein SAR1 (InterPro:IPR006687), ARF/SAR superfamily (InterPro:IPR006689); BEST Arabidopsis thaliana protein match is: ATSAR2 (ARABIDOPSIS THALIANA SECRETION-ASSOCIATED RAS SUPER FAMILY 2); GTP binding (TAIR:AT4G02080.1); Has 5231 Blast hits to 5228 proteins in 291 species: Archae - 2; Bacteria - 39; Metazoa - 2741; Fungi - 852; Plants - 681; Viruses - 0; Other Eukaryotes - 916 (source: NCBI BLink). |
| Gma.4531.1.A1_a_at | 0.058 | 1.96622 | Glyma06g10710.1 | AT4G08950.1 | EXORDIUM (EXO); FUNCTIONS IN: molecular_function unknown; INVOLVED IN: response to brassinosteroid stimulus; LOCATED IN: cell wall, plant-type cell wall; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Phosphate-induced protein 1 conserved region (InterPro:IPR006766); BEST Arabidopsis thaliana protein match is: PHI-1 (PHOSPHATE-INDUCED 1) (TAIR:AT1G35140.1); Has 234 Blast hits to 233 proteins in 17 species: Archae - 0; Bacteria - 2; Metazoa - 0; Fungi - 0; Plants - 232; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| GmaAffx.83336.1.S1_at | 0.0634 | 1.96165 | Glyma18g50860.1 | AT5G28680.1 | Receptor-like kinase required for maintenance of pollen tube growth. Display polar localization at the plasma membrane of the pollen tube tip. |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.67160.1.S1_at | 0.0416 | 1.92362 | Glyma13g24070.1 | AT2G33990.1 | IQ-domain 9 (iqd9); FUNCTIONS IN: calmodulin binding; INVOLVED IN: biological_process unknown; LOCATED IN: plasma membrane; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: IQ calmodulin-binding region (InterPro:IPR000048); BEST Arabidopsis thaliana protein match is: IQD10 (IQ-domain 10); calmodulin binding (TAIR:AT3G15050.1); Has 354 Blast hits to 353 proteins in 14 species: Archae - 0; Bacteria - 0; Metazoa - 1; Fungi - 0; Plants - 348; Viruses - 0; Other Eukaryotes - 5 (source: NCBI BLink). |
| Gma.8111.1.S1_at | 0.0338 | 1.92163 | Glyma08g42850.1 | AT3G20410.1 | calmodulin-domain protein kinase CDPK isoform 9 (CPK9) |
| Gma.2457.1.S1_at | 0.0317 | 1.92013 | No Soybean Match Identified, E<10E-30, Perecent ID>95% | AT1G53050.1 | protein kinase family protein; FUNCTIONS IN: protein serine/threonine kinase activity, protein kinase activity, kinase activity, ATP binding; INVOLVED IN: protein amino acid phosphorylation, N-terminal protein myristoylation; LOCATED IN: plasma membrane; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Protein kinase, ATP binding site (InterPro:IPR017441), Protein kinase, core (InterPro:IPR000719), Serine/threonine protein kinase (InterPro:IPR002290), Serine/threonine protein kinase-related (InterPro:IPR017442), Protein kinase-like (InterPro:IPR011009), Serine/threonine protein kinase, active site (InterPro:IPR008271); BEST Arabidopsis thaliana protein match is: protein kinase family protein (TAIR:AT1G09600.1); Has 91001 Blast hits to 89920 proteins in 2875 species: Archae - 45; Bacteria - 7929; Metazoa - 39540; Fungi - 8389; Plants - 17795; Viruses - 415; Other Eukaryotes - 16888 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.3990.1.S1_s_at | 0.0406 | 1.90461 | Glyma08g14370.1 | AT5G64220.2 | calmodulin-binding protein; FUNCTIONS IN: transcription regulator activity, calmodulin binding; INVOLVED IN: biological_process unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: CG-1 (InterPro:IPR005559), IQ calmodulin-binding region (InterPro:IPR000048), Ankyrin (InterPro:IPR002110); BEST Arabidopsis thaliana protein match is: EICBP.B (ETHYLENE INDUCED CALMODULIN BINDING PROTEIN); calmodulin binding / transcription activator/ transcription regulator (TAIR:AT5G09410.2); Has 12371 Blast hits to 7464 proteins in 391 species: Archae - 23; Bacteria - 619; Metazoa - 7363; Fungi - 661; Plants - 505; Viruses - 67; Other Eukaryotes - 3133 (source: NCBI BLink). |
| Gma.10599.1.S1_at | 0.0962 | 1.89496 | Glyma20g26920.1 | AT2G14820.1 | A member of the NPY gene family (NPY1/AT4G31820, NPY2/AT2G14820, NPY3/AT5G67440, NPY4/AT2G23050, NPY5/AT4G37590). Involved in auxin-mediated organogenesis. |
| GmaAffx.63840.1.S1_at | 0.0499 | 1.8749 | No Soybean Match Identified, E<10E-30, Perecent ID>95% | AT5G51260.1 | acid phosphatase, putative; FUNCTIONS IN: acid phosphatase activity; INVOLVED IN: biological_process unknown; LOCATED IN: endomembrane system; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Acid phosphatase (Class B) (InterPro:IPR005519), Vegetative storage protein/acid phosphatase (InterPro:IPR014403), Acid phosphatase, plant (InterPro:IPR010028); BEST Arabidopsis thaliana protein match is: acid phosphatase, putative (TAIR:AT4G25150.1); Has 538 Blast hits to 538 proteins in 152 species: Archae - 0; Bacteria - 252; Metazoa - 2; Fungi - 0; Plants - 237; Viruses - 0; Other Eukaryotes - 47 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.90671.1.S1_s_at | 0.0448 | 1.8148 | Glyma12g33550.1 | AT5G60860.1 | Arabidopsis Rab GTPase homolog A1f (AtRABA1f); FUNCTIONS IN: GTP binding; INVOLVED IN: protein transport, small GTPase mediated signal transduction; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Ras GTPase (InterPro:IPR001806), Small GTP-binding protein (InterPro:IPR005225), Ras (InterPro:IPR013753), Ras small GTPase, Rab type (InterPro:IPR003579), Rab11-related (InterPro:IPR015595); BEST Arabidopsis thaliana protein match is: AtRABA1g (Arabidopsis Rab GTPase homolog A1g); GTP binding (TAIR:AT3G15060.1); Has 21819 Blast hits to 21779 proteins in 607 species: Archae - 23; Bacteria - 103; Metazoa - 12072; Fungi - 2821; Plants - 1861; Viruses - 19; Other Eukaryotes - 4920 (source: NCBI BLink). |
| GmaAffx.52514.1.S1_at | 0.0812 | 1.7863 | Glyma20g00790.1 | AT3G11920.1 | glutaredoxin-related; FUNCTIONS IN: electron carrier activity, protein disulfide oxidoreductase activity; INVOLVED IN: intracellular signaling cascade, cell redox homeostasis |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.91247.1.S1_at | 0.0274 | 1.78081 | Glyma15g09490.1 | AT4G18950.1 | ankyrin protein kinase, putative; FUNCTIONS IN: protein serine/threonine/tyrosine kinase activity, kinase activity; INVOLVED IN: regulation of signal transduction, protein amino acid phosphorylation; LOCATED IN: nucleus, cytoplasm; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Protein kinase, core (InterPro:IPR000719), Integrin-linked protein kinase (InterPro:IPR016253), Tyrosine protein kinase (InterPro:IPR001245), Serine/threonine protein kinase-related (InterPro:IPR017442), Protein kinase-like (InterPro:IPR011009), Ankyrin (InterPro:IPR002110); BEST Arabidopsis thaliana protein match is: ankyrin protein kinase, putative (TAIR:AT3G58760.1); Has 116131 Blast hits to 99643 proteins in 3446 species: Archae - 96; Bacteria - 9447; Metazoa - 52988; Fungi - 8496; Plants - 19258; Viruses - 634; Other Eukaryotes - 25212 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.16954.2.A1_at | 0.056 | 1.74475 | Glyma18g03100.1 | AT2G40116.1 | phosphoinositide-specific phospholipase C family protein; FUNCTIONS IN: phospholipase C activity, phosphoinositide phospholipase C activity, phosphoric diester hydrolase activity; INVOLVED IN: signal transduction, intracellular signaling cascade, lipid metabolic process; LOCATED IN: cellular_component unknown; CONTAINS InterPro DOMAIN/s: Phospholipase C, phosphoinositol-specific, EF-hand-like (InterPro:IPR015359), Phospholipase C, phosphatidylinositol-specific , X region (InterPro:IPR000909), PLC-like phosphodiesterase, TIM beta/alpha-barrel domain (InterPro:IPR017946), C2 membrane targeting protein (InterPro:IPR018029), C2 calcium/lipid-binding region, CaLB (InterPro:IPR008973), Phospholipase C, phosphoinositol-specific, C-terminal (PLC) (InterPro:IPR001192), C2 calcium-dependent membrane targeting (InterPro:IPR000008), Phospholipase C, phosphatidylinositol-specific, Y domain (InterPro:IPR001711); BEST Arabidopsis thaliana protein match is: ATPLC2 (PHOSPHOLIPASE C 2); phospholipase C (TAIR:AT3G08510.2); Has 2150 Blast hits to 1783 proteins in 218 species: Archae - 0; Bacteria - 0; Metazoa - 1549; Fungi - 2 |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.40428.1.S1_at | 0.04 | 1.72827 | Glyma08g17650.1 | AT1G79570.1 | protein kinase family protein; FUNCTIONS IN: protein serine/threonine/tyrosine kinase activity, protein kinase activity; INVOLVED IN: protein amino acid phosphorylation; LOCATED IN: cytosol; EXPRESSED IN: male gametophyte, pollen tube; EXPRESSED DURING: L mature pollen stage, M germinated pollen stage; CONTAINS InterPro DOMAIN/s: Octicosapeptide/Phox/Bem1p (InterPro:IPR000270), Protein kinase, ATP binding site (InterPro:IPR017441), Protein kinase, core (InterPro:IPR000719), Tyrosine protein kinase (InterPro:IPR001245), Protein kinase-like (InterPro:IPR011009); BEST Arabidopsis thaliana protein match is: protein kinase family protein (TAIR:AT1G16270.1); Has 86578 Blast hits to 85332 proteins in 3442 species: Archae - 47; Bacteria - 6840; Metazoa - 39346; Fungi - 6791; Plants - 18389; Viruses - 433; Other Eukaryotes - 14732 (source: NCBI BLink). |
| GmaAffx.7365.1.S1_at | 0.0917 | 1.7188 | Glyma07g21110.1 | AT2G23460.1 | encodes a novel G-alpha protein that shares similarity to plant, yeast, and animal G-alpha proteins at the C-terminus. |

| | | | | | |
|---|---|---|---|---|---|
| Gma.13604.1.S1_at | 0.0514 | 1.71601 | Glyma14g06420.3 | AT2G40120.1 | protein kinase family protein; FUNCTIONS IN: protein serine/threonine kinase activity, protein kinase activity, kinase activity, ATP binding; INVOLVED IN: protein amino acid phosphorylation; LOCATED IN: cytoplasm; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Protein kinase, core (InterPro:IPR000719), Serine/threonine protein kinase (InterPro:IPR002290), Serine/threonine protein kinase-related (InterPro:IPR017442), Protein kinase-like (InterPro:IPR011009), Serine/threonine protein kinase, active site (InterPro:IPR008271); BEST Arabidopsis thaliana protein match is: protein kinase family protein (TAIR:AT3G17750.1); Has 65627 Blast hits to 64489 proteins in 2041 species: Archae - 50; Bacteria - 5400; Metazoa - 28985; Fungi - 7798; Plants - 8456; Viruses - 355; Other Eukaryotes - 14583 (source: NCBI BLink). |
| Gma.4264.1.S1_at | 0.0775 | 1.57981 | Glyma08g29060.1 | AT3G02750.2 | protein phosphatase 2C family protein / PP2C family protein; FUNCTIONS IN: protein serine/threonine phosphatase activity, catalytic activity; LOCATED IN: chloroplast; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Protein phosphatase 2C-related (InterPro:IPR001932), Protein phosphatase 2C (InterPro:IPR015655), Protein phosphatase 2C, N-terminal (InterPro:IPR014045); BEST Arabidopsis thaliana protein match is: protein phosphatase 2C, putative / PP2C, putative (TAIR:AT5G36250.1); Has 3763 Blast hits to 3747 proteins in 227 species: Archae - 0; Bacteria - 10; Metazoa - 1174; Fungi - 424; Plants - 1253; Viruses - 5; Other Eukaryotes - 897 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.33420.1.A1_at | 0.0514 | 1.51923 | Glyma10g04410.3 | AT2G20470.1 | kinase; FUNCTIONS IN: kinase activity; INVOLVED IN: protein amino acid phosphorylation; EXPRESSED IN: male gametophyte; EXPRESSED DURING: M germinated pollen stage; CONTAINS InterPro DOMAIN/s: Protein kinase, ATP binding site (InterPro:IPR017441), Serine/threonine protein kinase (InterPro:IPR002290), Serine/threonine protein kinase-related (InterPro:IPR017442), Protein kinase, C-terminal (InterPro:IPR017892), Protein kinase-like (InterPro:IPR011009), Serine/threonine protein kinase, active site (InterPro:IPR008271), AGC-kinase, C-terminal (InterPro:IPR000961), Protein kinase, core (InterPro:IPR000719); BEST Arabidopsis thaliana protein match is: protein kinase, putative (TAIR:AT1G03920.1); Has 80529 Blast hits to 79531 proteins in 2762 species: Archae - 65; Bacteria - 7647; Metazoa - 33811; Fungi - 7913; Plants - 14736; Viruses - 399; Other Eukaryotes - 15958 (source: NCBI BLink). |
| Gma.15570.1.S1_at | 0.0557 | 1.50714 | Glyma18g17580.1 | AT1G13750.1 | calcineurin-like phosphoesterase family protein |
| Gma.7177.2.A1_a_at | 0.043 | 1.49422 | Glyma16g10820.2 | AT4G13020.1 | Encodes a member of the cdc2+ family of protein kinases MHK. Similar to the mak genes of rats. mak encodes a protein kinase that may play a role in spermatogenesis. |
| GmaAffx.19817.1.A1_at | 0.0953 | 1.48291 | Glyma13g38800.2 | AT3G22190.1 | IQ-domain 5 (IQD5); FUNCTIONS IN: calmodulin binding |
| Gma.2520.1.S1_at | 0.0996 | 1.47658 | No Soybean Match Identified, E<10E-30, Perecent ID>95% | AT2G46500.2 | Phosphoinositide kinase which undergo autophosphorylation and phosphorylate serine/threonine residues of protein substrates. Contains phosphoinositide 3/4-kinase and ubiquitin-like domains. Phosphorylates PUFD1 and RPN10 in vitro. |

| | | | | | |
|---|---|---|---|---|---|
| Gma.2356.1.A1_at | 0.094 | 1.43125 | Glyma13g10480.1 | AT5G41990.1 | Encodes a member of the WNK family (9 members in all) of protein kinases, the structural design of which is clearly distinct from those of other known protein kinases, such as receptor-like kinases and mitogen-activated protein kinases. Interacts specifically with and phosphorylates AtVHA-C, subunit C of the vacuolar H+-ATPase. |
| Gma.9593.1.A1_at | 0.0724 | 1.42556 | Glyma08g12050.1 | AT4G08810.1 | Calcium binding protein involved in cryptochrome and phytochrome coaction |
| Gma.5188.1.S1_at | 0.0996 | 1.41923 | Glyma08g23550.2 | AT4G31860.1 | protein phosphatase 2C, putative / PP2C, putative; FUNCTIONS IN: protein serine/threonine phosphatase activity, catalytic activity; INVOLVED IN: protein amino acid dephosphorylation, N-terminal protein myristoylation; LOCATED IN: plasma membrane; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Protein phosphatase 2C, manganese/magnesium aspartate binding site (InterPro:IPR000222), Protein phosphatase 2C-related (InterPro:IPR001932), Protein phosphatase 2C (InterPro:IPR015655), Protein phosphatase 2C, N-terminal (InterPro:IPR014045); BEST Arabidopsis thaliana protein match is: protein phosphatase 2C, putative / PP2C, putative (TAIR:AT2G25070.1); Has 4492 Blast hits to 4328 proteins in 297 species: Archae - 2; Bacteria - 128; Metazoa - 1456; Fungi - 538; Plants - 1338; Viruses - 7; Other Eukaryotes - 1023 (source: NCBI BLink). |
| Gma.12559.1.A1_at | 0.0053 | 1.36355 | Glyma10g12110.1 | AT3G12160.1 | Encodes RABA4D, a member of the Arabidopsis RabA4 subfamily of Rab GTPase proteins. It is transported in exocytic vesicles to the apical tip of pollen tubes where it appears to promote tip growth. Proper localization of RabA4d depends on ROP1, RIC3, and RIC4 activity. |
| GmaAffx.34228.1.A1_at | 0.094 | 1.35971 | Glyma20g27070.2 | AT5G67560.1 | A member of ARF-like GTPase family. |

| | | | | | |
|---|---|---|---|---|---|
| Gma.5980.1.S1_at | 0.0685 | 1.35101 | Ambiguous Hit | AT3G17700.1 | cyclic nucleotide-binding transporter 1, member of a family of cyclic nucleotide gated channels. |
| GmaAffx.22165.1.S1_at | 0.0771 | 1.34317 | Glyma17g09680.1 | AT5G16760.1 | Encodes a inositol 1,3,4-trisphosphate 5/6-kinase. |
| GmaAffx.53957.1.S1_at | 0.0337 | 1.2999 | Glyma17g15550.2 | AT4G17530.1 | ATRAB1C; FUNCTIONS IN: GTP binding; INVOLVED IN: protein transport, small GTPase mediated signal transduction; LOCATED IN: plasma membrane, vacuole, membrane; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Ras GTPase (InterPro:IPR001806), Small GTP-binding protein (InterPro:IPR005225), Ras (InterPro:IPR013753), Ras small GTPase, Rab type (InterPro:IPR003579); BEST Arabidopsis thaliana protein match is: ATRAB1A; GTP binding (TAIR:AT5G47200.1); Has 23593 Blast hits to 23542 proteins in 647 species: Archae - 17; Bacteria - 111; Metazoa - 13220; Fungi - 2827; Plants - 2160; Viruses - 19; Other Eukaryotes - 5239 (source: NCBI BLink). |
| Gma.10801.1.S1_at | 0.0685 | 1.25096 | Glyma03g15850.1 | AT3G54020.1 | Inositol phosphorylceramide synthase |
| Gma.596.1.S1_at | 0.0268 | -2.9551 | Glyma13g34580.4 | AT2G42590.2 | 14-3-3 gene. Binds calcium and displays induced structural changes. |

| | | | | | |
|---|---|---|---|---|---|
| Gma.3185.1.S1_at | 0.0539 | -2.08403 | Glyma18g16150.1 | AT1G14390.1 | leucine-rich repeat transmembrane protein kinase, putative; FUNCTIONS IN: protein binding, protein serine/threonine kinase activity, protein tyrosine kinase activity, protein kinase activity, ATP binding; INVOLVED IN: transmembrane receptor protein tyrosine kinase signaling pathway, protein amino acid phosphorylation; LOCATED IN: endomembrane system; EXPRESSED IN: inflorescence meristem, flower, root; EXPRESSED DURING: petal differentiation and expansion stage; CONTAINS InterPro DOMAIN/s: Protein kinase, core (InterPro:IPR000719), Serine/threonine protein kinase (InterPro:IPR002290), Leucine-rich repeat (InterPro:IPR001611), Tyrosine protein kinase (InterPro:IPR001245), Protein kinase-like (InterPro:IPR011009); BEST Arabidopsis thaliana protein match is: leucine-rich repeat transmembrane protein kinase, putative (TAIR:AT2G02780.1); Has 60751 Blast hits to 33447 proteins in 1087 species: Archae - 20; Bacteria - 3500; Metazoa - 15937; Fungi - 887; Plants - 36384; Viruses - 79; Other Eukaryotes - 3944 (source: NCBI BLink). |
| Gma.16189.2.S1_at | 0.058 | -1.98407 | Glyma13g34990.1 | AT3G15260.2 | protein phosphatase 2C, putative / PP2C, putative; FUNCTIONS IN: protein serine/threonine phosphatase activity, catalytic activity; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Protein phosphatase 2C-related (InterPro:IPR001932), Protein phosphatase 2C (InterPro:IPR015655), Protein phosphatase 2C, N-terminal (InterPro:IPR014045); BEST Arabidopsis thaliana protein match is: protein phosphatase 2C, putative / PP2C, putative (TAIR:AT4G28400.1); Has 4540 Blast hits to 4534 proteins in 402 species: Archae - 3; Bacteria - 375; Metazoa - 1354; Fungi - 535; Plants - 1305; Viruses - 11; Other Eukaryotes - 957 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.35281.1.S1_at | 0.0782 | -1.94904 | Glyma04g06260.1 | AT5G24970.1 | ABC1 family protein; FUNCTIONS IN: protein kinase activity, ATP binding; INVOLVED IN: protein amino acid phosphorylation; LOCATED IN: cellular_component unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: ABC-1 (InterPro:IPR004147), Protein kinase, core (InterPro:IPR000719), Protein kinase-like (InterPro:IPR011009); BEST Arabidopsis thaliana protein match is: ABC1 family protein (TAIR:AT1G79600.1); Has 7448 Blast hits to 7430 proteins in 1103 species: Archae - 69; Bacteria - 2644; Metazoa - 345; Fungi - 303; Plants - 352; Viruses - 14; Other Eukaryotes - 3721 (source: NCBI BLink). |
| GmaAffx.35378.1.S1_s_at | 0.0947 | -1.88159 | Glyma07g01350.1 | AT3G13690.1 | protein kinase family protein; FUNCTIONS IN: protein tyrosine kinase activity, protein kinase activity, kinase activity, ATP binding; INVOLVED IN: protein amino acid phosphorylation; LOCATED IN: cellular_component unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Protein kinase, core (InterPro:IPR000719), Tyrosine protein kinase, active site (InterPro:IPR008266), Serine/threonine protein kinase-related (InterPro:IPR017442), Protein kinase-like (InterPro:IPR011009); BEST Arabidopsis thaliana protein match is: protein kinase family protein (TAIR:AT1G55200.1); Has 81404 Blast hits to 80467 proteins in 3236 species: Archae - 44; Bacteria - 7601; Metazoa - 35853; Fungi - 6033; Plants - 18144; Viruses - 313; Other Eukaryotes - 13416 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.37155.1.S1_at | 0.063 | -1.81305 | Glyma13g30830.1 | AT1G28440.1 | HAESA-Like 1 (HSL1); FUNCTIONS IN: protein serine/threonine kinase activity, kinase activity, ATP binding; INVOLVED IN: transmembrane receptor protein tyrosine kinase signaling pathway, protein amino acid phosphorylation; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Protein kinase, ATP binding site (InterPro:IPR017441), Protein kinase, core (InterPro:IPR000719), Leucine-rich repeat, N-terminal (InterPro:IPR013210), Leucine-rich repeat (InterPro:IPR001611), Serine/threonine protein kinase-related (InterPro:IPR017442), Protein kinase-like (InterPro:IPR011009), Serine/threonine protein kinase, active site (InterPro:IPR008271); BEST Arabidopsis thaliana protein match is: HAE (HAESA); ATP binding / kinase/ protein kinase/ protein serine/threonine kinase (TAIR:AT4G28490.1); Has 146860 Blast hits to 94798 proteins in 2852 species: Archae - 79; Bacteria - 12502; Metazoa - 57487; Fungi - 6803; Plants - 50326; Viruses - 373; Other Eukaryotes - 19290 (source: NCBI BLink). |
| Gma.8363.1.S1_at | 0.0873 | -1.68119 | Glyma05g22380.1 | AT2G14820.1 | A member of the NPY gene family (NPY1/AT4G31820, NPY2/AT2G14820, NPY3/AT5G67440, NPY4/AT2G23050, NPY5/AT4G37590). Involved in auxin-mediated organogenesis. |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.75246.1.S1_at | 0.0679 | -1.62157 | Glyma02g04150.2 | AT1G60800.1 | NSP-INTERACTING KINASE 3 (NIK3); FUNCTIONS IN: kinase activity; INVOLVED IN: protein amino acid phosphorylation; LOCATED IN: plasma membrane; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Protein kinase, ATP binding site (InterPro:IPR017441), Protein kinase, core (InterPro:IPR000719), Leucine-rich repeat, N-terminal (InterPro:IPR013210), Leucine-rich repeat (InterPro:IPR001611), Serine/threonine protein kinase-related (InterPro:IPR017442), Protein kinase-like (InterPro:IPR011009), Serine/threonine protein kinase, active site (InterPro:IPR008271); BEST Arabidopsis thaliana protein match is: NIK2 (NSP-INTERACTING KINASE 2); ATP binding / protein binding / protein kinase/ protein serine/threonine kinase (TAIR:AT3G25560.1); Has 119774 Blast hits to 86788 proteins in 3214 species: Archae - 64; Bacteria - 8965; Metazoa - 39111; Fungi - 5992; Plants - 49549; Viruses - 384; Other Eukaryotes - 15709 (source: NCBI BLink). |
| GmaAffx.12237.1.S1_at | 0.0724 | -1.56592 | Glyma06g20350.1 | AT1G74690.1 | IQ-domain 31 (IQD31); FUNCTIONS IN: calmodulin binding; INVOLVED IN: biological_process unknown; LOCATED IN: cytosol, plasma membrane; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: IQ calmodulin-binding region (InterPro:IPR000048); BEST Arabidopsis thaliana protein match is: IQD30; calmodulin binding (TAIR:AT1G18840.2); Has 7267 Blast hits to 5232 proteins in 414 species: Archae - 15; Bacteria - 567; Metazoa - 3152; Fungi - 664; Plants - 611; Viruses - 38; Other Eukaryotes - 2220 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.9861.1.S1_at | 0.0832 | -1.52655 | Glyma15g07820.2 | AT3G09010.1 | protein kinase family protein; FUNCTIONS IN: protein serine/threonine kinase activity, protein kinase activity, kinase activity, ATP binding; INVOLVED IN: protein amino acid phosphorylation; LOCATED IN: cellular_component unknown; EXPRESSED IN: 20 plant structures; EXPRESSED DURING: 12 growth stages; CONTAINS InterPro DOMAIN/s: Protein kinase, ATP binding site (InterPro:IPR017441), Protein kinase, core (InterPro:IPR000719), Serine/threonine protein kinase-related (InterPro:IPR017442), Protein kinase-like (InterPro:IPR011009), Serine/threonine protein kinase, active site (InterPro:IPR008271); BEST Arabidopsis thaliana protein match is: protein kinase family protein (TAIR:AT1G16670.1); Has 85198 Blast hits to 84212 proteins in 3264 species: Archae - 56; Bacteria - 7432; Metazoa - 37605; Fungi - 6636; Plants - 18795; Viruses - 359; Other Eukaryotes - 14315 (source: NCBI BLink). |
| GmaAffx.53682.1.S1_at | 0.0882 | -1.47914 | Glyma15g08350.1 | AT1G28340.1 | Receptor Like Protein 4 (AtRLP4); FUNCTIONS IN: protein binding; INVOLVED IN: signal transduction |
| Gma.17770.1.S1_at | 0.0911 | -1.39983 | Glyma09g34940.3 | AT1G31420.1 | Encodes a plasma membrane localized leucine-rich repeat receptor kinase that is involved in cell wall elongation. Loss of function mutations of FEI1 and FEI2 exhibit defects in root and hypocotyl cell elongation. Double mutants are defective in cell wall biosynthesis and have thick hypocotyls, and short, thick roots. |
| GmaAffx.18412.1.S1_at | 0.0418 | -1.37105 | Glyma18g51100.1 | AT2G03890.1 | Phosphoinositide kinase which undergo autophosphorylation and phosphorylate serine/threonine residues of protein substrates. Contains phosphoinositide 3/4-kinase and ubiquitin-like domains. |
| Gma.1814.1.S1_at | 0.0481 | -1.33354 | Glyma12g28650.6 | AT3G11730.1 | Encodes a member of the Rab GTPase family of proteins. This protein interacts with the tail region of a myosin XI protein (AT5G43900) in a GTP-dependent manner. It has also been identified as an isoprenylated protein. |

| | | | | | |
|---|---|---|---|---|---|
| Gma.7541.1.S1_at | 0.0802 | -1.27808 | No Soybean Match Identified, E<10E-30, Perecent ID>95% | AT4G06599.1 | ubiquitin family protein; FUNCTIONS IN: phosphoprotein phosphatase activity; INVOLVED IN: protein modification process; LOCATED IN: nucleus; CONTAINS InterPro DOMAIN/s: Ubiquitin (InterPro:IPR000626), HAD-superfamily hydrolase, subfamily IIID (InterPro:IPR011943), NLI interacting factor (InterPro:IPR004274); Has 297 Blast hits to 297 proteins in 94 species: Archae - 0; Bacteria - 4; Metazoa - 118; Fungi - 37; Plants - 59; Viruses - 6; Other Eukaryotes - 73 (source: NCBI BLink). |
| Gma.6928.1.S1_at | 0.0966 | -1.22639 | Glyma11g14240.1 | AT5G59010.1 | protein kinase-related; FUNCTIONS IN: protein tyrosine kinase activity, binding, protein kinase activity, ATP binding; INVOLVED IN: protein amino acid phosphorylation, N-terminal protein myristoylation; LOCATED IN: plasma membrane; EXPRESSED IN: 20 plant structures; EXPRESSED DURING: 10 growth stages; CONTAINS InterPro DOMAIN/s: Tetratricopeptide-like helical (InterPro:IPR011990), Protein kinase, core (InterPro:IPR000719), Tyrosine protein kinase (InterPro:IPR001245), Protein kinase-like (InterPro:IPR011009); BEST Arabidopsis thaliana protein match is: protein kinase family protein (TAIR:AT3G54030.1); Has 21265 Blast hits to 20899 proteins in 886 species: Archae - 6; Bacteria - 1353; Metazoa - 6281; Fungi - 197; Plants - 12179; Viruses - 98; Other Eukaryotes - 1151 (source: NCBI BLink). |

| Affy ID | Q-value | Fold change | Soybean Gene | Arabidopsis homolog | Description |
|---|---|---|---|---|---|
| Gma.7623.1.A1_at | 0.03126 | 87.8905 | Glyma07g06750.1 | AT2G46240.1 | A member of Arabidopsis BAG (Bcl-2-associated athanogene) proteins, plant homologs of mammalian regulators of apoptosis. Expression of BAG6 in leaves was strongly induced by heat stress. Knockout mutants exhibited enhanced susceptibility to fungal pathogen Botrytis cinerea. Plant BAG proteins are multi-functional and remarkably similar to their animal counterparts, as they regulate apoptotic-like processes ranging from pathogen attack, to abiotic stress, to plant development. |
| GmaAffx.75438.1.S1_ | 0.06412 | 62.3066 | Glyma17g13820.1 | AT1G76650.1 | CALMODULIN-LIKE 38 (CML38); FUNCTIONS IN: calcium ion binding; INVOLVED IN: response to wounding; LOCATED IN: plasma membrane; EXPRESSED IN: 19 plant structures; EXPRESSED DURING: 12 growth stages; CONTAINS InterPro DOMAIN/s: EF-HAND 1 (InterPro:IPR018247), EF-HAND 2 (InterPro:IPR018249), Calcium-binding EF-hand (InterPro:IPR002048), EF-Hand type (InterPro:IPR011992), EF hand (InterPro:IPR018248); BEST Arabidopsis thaliana protein match is: calmodulin-related protein, putative (TAIR:AT1G76640.1); Has 11212 Blast hits to 8813 proteins in 1020 species: Archae - 0; Bacteria - 34; Metazoa - 5021; Fungi - 1984; Plants - 2417; Viruses - 0; Other Eukaryotes - 1756 (source: NCBI BLink). |
| GmaAffx.25859.1.S1_ | 0.07641 | 30.2139 | Glyma14g06910.1 | AT3G46230.1 | member of the class I small heat-shock protein (sHSP) family, which accounts for the majority of sHSPs in maturing seeds |
| GmaAffx.39393.1.S1_ | 0.03473 | 16.1619 | Glyma10g32000.1 | AT4G10250.1 | Columbia endomembrane-localized small heat shock protein |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.91194.1.S1 | 0.03761 | 14.4989 | Glyma01g41920.1 | AT4G17260.1 | L-lactate dehydrogenase, putative; FUNCTIONS IN: in 6 functions; INVOLVED IN: response to salt stress, response to abscisic acid stimulus; LOCATED IN: plasma membrane; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: L-lactate dehydrogenase (InterPro:IPR011304), NAD(P)-binding (InterPro:IPR016040), L-lactate/malate dehydrogenase (InterPro:IPR001557), Lactate/malate dehydrogenase (InterPro:IPR001236), L-lactate dehydrogenase, active site (InterPro:IPR018177), Lactate dehydrogenase/glycoside hydrolase, family 4, C-terminal (InterPro:IPR015955); BEST Arabidopsis thaliana protein match is: PMDH2 (peroxisomal NAD-malate dehydrogenase 2); malate dehydrogenase (TAIR:AT5G09660.2); Has 9812 Blast hits to 9805 proteins in 2184 species: Archae - 185; Bacteria - 5417; Metazoa - 1010; Fungi - 354; Plants - 275; Viruses - 0; Other Eukaryotes - 2571 (source: NCBI BLink). |
| Gma.17947.1.S1_at | 0.05654 | 13.4806 | Glyma08g07340.1 | AT2G29500.1 | 17.6 kDa class I small heat shock protein (HSP17.6B-CI); FUNCTIONS IN: molecular_function unknown; INVOLVED IN: response to oxidative stress, response to cyclopentenone, response to high light intensity, response to hydrogen peroxide, response to heat; EXPRESSED IN: 17 plant structures; EXPRESSED DURING: 9 growth stages; CONTAINS InterPro DOMAIN/s: Heat shock protein Hsp20 (InterPro:IPR002068), HSP20-like chaperone (InterPro:IPR008978); BEST Arabidopsis thaliana protein match is: 17.8 kDa class I heat shock protein (HSP17.8-CI) (TAIR:AT1G07400.1); Has 4233 Blast hits to 4233 proteins in 935 species: Archae - 103; Bacteria - 2307; Metazoa - 65; Fungi - 220; Plants - 1015; Viruses - 0; Other Eukaryotes - 523 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.14272.1.S1_at | 0.03956 | 11.6943 | Glyma10g28300.1 | AT3G05890.1 | RARE-COLD-INDUCIBLE 2B (RCI2B); INVOLVED IN: response to cold; LOCATED IN: integral to membrane; EXPRESSED IN: 19 plant structures; EXPRESSED DURING: 10 growth stages; CONTAINS InterPro DOMAIN/s: Uncharacterised protein family UPF0057 (InterPro:IPR000612); BEST Arabidopsis thaliana protein match is: RCI2A (RARE-COLD-INDUCIBLE 2A) (TAIR:AT3G05880.1); Has 905 Blast hits to 905 proteins in 327 species: Archae - 0; Bacteria - 464; Metazoa - 38; Fungi - 185; Plants - 185; Viruses - 0; Other Eukaryotes - 33 (source: NCBI BLink). |
| Gma.10763.1.S1_at | 0.06636 | 11.3641 | Glyma04g05720.1 | AT5G12020.1 | 17.6 KDA CLASS II HEAT SHOCK PROTEIN (HSP17.6II); INVOLVED IN: response to heat; LOCATED IN: cellular_component unknown; EXPRESSED IN: 9 plant structures; EXPRESSED DURING: M germinated pollen stage, 4 anthesis, C globular stage, petal differentiation and expansion stage; CONTAINS InterPro DOMAIN/s: Heat shock protein Hsp20 (InterPro:IPR002068), HSP20-like chaperone (InterPro:IPR008978); BEST Arabidopsis thaliana protein match is: AT-HSP17.6A (ARABIDOPSIS THALIANA HEAT SHOCK PROTEIN 17.6A); unfolded protein binding (TAIR:AT5G12030.1); Has 4227 Blast hits to 4227 proteins in 942 species: Archae - 102; Bacteria - 2394; Metazoa - 83; Fungi - 183; Plants - 932; Viruses - 0; Other Eukaryotes - 533 (source: NCBI BLink). |
| GmaAffx.29929.1.S1_ | 0.02679 | 11.2499 | Glyma20g29410.1 | AT4G25480.1 | encodes a member of the DREB subfamily A-1 of ERF/AP2 transcription factor family (CBF3). The protein contains one AP2 domain. There are six members in this subfamily, including CBF1, CBF2, and CBF3. This gene is involved in response to low temperature and abscisic acid. |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.2044.1.S1_at | 0.07224 | 10.7706 | Glyma09g31740.1 | AT5G66400.2 | | Belongs to the dehydrin protein family, which contains highly conserved stretches of 7-17 residues that are repetitively scattered in their sequences, the K-, S-, Y- and lysine rich segments. ABA- and drought-induced glycine-rice dehydrin protein. The ABA-induced expression of RAB18 was reduced following ACC application, indicating that ethylene inhibits the ABA signaling pathway. RAB18 is also expressed in response to the formation of the phospholipid diacylglycerol pyrophosphate. COR47 and RAB18 double overexpressor plants are cold tolerant. Expressed in guard cells. |
| GmaAffx.68621.1.A1 | 0.09884 | 9.43873 | Glyma20g29770.1 | AT5G52300.2 | | encodes a protein that is induced in expression in response to water deprivation such as cold, high-salt, and dessication. The response appears to be via abscisic acid. The promoter region contains two ABA-responsive elements (ABREs) that are required for the dehydration-responsive expression of rd29B as cis-acting elements. Protein is a member of a gene family with other members found plants, animals and fungi. |
| Gma.17802.1.S1_at | 0.02679 | 8.63685 | Glyma04g34890.1 | AT1G19020.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: response to oxidative stress; LOCATED IN: cellular_component unknown; EXPRESSED IN: 19 plant structures; EXPRESSED DURING: 9 growth stages; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT3G48180.1); Has 58 Blast hits to 58 proteins in 9 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 58; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| GmaAffx.88182.1.S1_ | 0.0448 | 8.42288 | Glyma03g35930.1 | AT2G35980.1 | | Encodes a protein whose sequence is similar to tobacco hairpin-induced gene (HIN1) and Arabidopsis non-race specific disease resistance gene (NDR1). Expression of this gene is induced by cucumber mosaic virus, spermine and during senescence. The gene product is localized to the chloroplast. |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.11004.1.S1_at | 0.03596 | 8.09389 | Glyma03g35920.1 | AT2G35980.1 | | Encodes a protein whose sequence is similar to tobacco hairpin-induced gene (HIN1) and Arabidopsis non-race specific disease resistance gene (NDR1). Expression of this gene is induced by cucumber mosaic virus, spermine and during senescence. The gene product is localized to the chloroplast. |
| Gma.9553.1.A1_at | 0.0317 | 8.0793 | Glyma14g06080.1 | AT2G40340.1 | | encodes a member of the DREB subfamily A-2 of ERF/AP2 transcription factor family. The protein contains one AP2 domain. There are eight members in this subfamily including DREB2A AND DREB2B that are involved in response to drought. |
| Gma.5637.1.S1_at | 0.05952 | 7.88493 | Glyma10g28610.1 | AT2G47180.1 | | Arabidopsis thaliana galactinol synthase 1 (AtGolS1); FUNCTIONS IN: transferase activity, transferring hexosyl groups, transferase activity, transferring glycosyl groups; INVOLVED IN: response to high light intensity, response to hydrogen peroxide, carbohydrate biosynthetic process, response to heat; LOCATED IN: cellular_component unknown; EXPRESSED IN: 19 plant structures; EXPRESSED DURING: 11 growth stages; CONTAINS InterPro DOMAIN/s: Glycosyl transferase, family 8 (InterPro:IPR002495); BEST Arabidopsis thaliana protein match is: AtGolS2 (Arabidopsis thaliana galactinol synthase 2); transferase, transferring glycosyl groups / transferase, transferring hexosyl groups (TAIR:AT1G56600.1); Has 818 Blast hits to 817 proteins in 194 species: Archae - 0; Bacteria - 56; Metazoa - 224; Fungi - 185; Plants - 235; Viruses - 68; Other Eukaryotes - 50 (source: NCBI BLink). |
| Gma.7526.1.A1_at | 0.03119 | 7.62069 | Glyma11g11430.1 | AT2G17840.1 | | Identified as drought-inducible gene by differential hybridization. Upregulated by high light, drought, cold and salt stress determined by microarray analysis. |
| GmaAffx.24571.1.S1_ | 0.02316 | 7.4588 | Glyma13g01150.1 | AT5G57560.1 | | Encodes a cell wall-modifying enzyme, rapidly upregulated in response to environmental stimuli |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.63418.1.S1_ | 0.0568 | 7.0813 | Glyma09g16810.1 | AT5G06720.1 | | peroxidase, putative; FUNCTIONS IN: electron carrier activity, peroxidase activity, heme binding; INVOLVED IN: response to oxidative stress; LOCATED IN: endomembrane system; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 11 growth stages; CONTAINS InterPro DOMAIN/s: Haem peroxidase (InterPro:IPR010255), Plant peroxidase (InterPro:IPR000823), Haem peroxidase, plant/fungal/bacterial (InterPro:IPR002016); BEST Arabidopsis thaliana protein match is: peroxidase, putative (TAIR:AT5G06730.1); Has 2941 Blast hits to 2925 proteins in 230 species: Archae - 0; Bacteria - 0; Metazoa - 2; Fungi - 109; Plants - 2780; Viruses - 2; Other Eukaryotes - 48 (source: NCBI BLink). |
| Gma.11195.1.S1_at | 0.06613 | 6.97681 | Glyma15g26590.1 | AT5G58070.1 | | Encodes a temperature-induced lipocalin TIL1. Involved in thermotolerance. Peripherally associated with plasma membrane. |
| GmaAffx.2469.1.S1_s | 0.03154 | 6.91654 | Glyma01g42500.2 | AT1G12610.1 | | Encodes a member of the DREB subfamily A-1 of ERF/AP2 transcription factor family (DDF1). The protein contains one AP2 domain. There are six members in this subfamily, including CBF1, CBF2, and CBF3. Overexpression of this gene results in delayed flowering and dwarfism, reduction of gibberellic acid biosynthesis, and increased tolerance to high levels of salt. This gene is expressed in all tissues examined, but most abundantly expressed in upper stems. Overexpression of this gene is also correlated with increased expression of GA biosynthetic genes and RD29A (a cold and drought responsive gene). Under salt stress it induces the expression of GAOX7, which encodes ad C20-GA inhibitor. |

| | | | | | |
|---|---|---|---|---|---|
| Gma.487.1.S1_at | 0.03369 | 6.85796 | Glyma18g43750.1 | AT3G11020.1 | encodes a member of the DREB subfamily A-2 of ERF/AP2 transcription factor family (DREB2B). The protein contains one AP2 domain. There are eight members in this subfamily including DREB2A. |
| GmaAffx.57556.1.S1_ | 0.02697 | 6.67398 | Glyma07g18550.2 | AT3G08970.1 | J domain protein localized in ER lumen. Can compensate for the growth defect in jem1 scj1 mutant yeast. Also shows similarity to HSP40 proteins and is induced by heat stress. At high temperatures, mutant alleles are not transmitted through the pollen due to defects in pollen tube growth. |
| Gma.2821.1.S1_at | 0.03623 | 6.40711 | Glyma01g42660.1 | AT4G11650.1 | osmotin-like protein |
| GmaAffx.30428.1.S1_ | 0.09965 | 6.37481 | Glyma17g08020.1 | AT1G16030.1 | heat shock protein 70B (Hsp70b); FUNCTIONS IN: ATP binding; INVOLVED IN: protein folding, response to heat; LOCATED IN: cytosol, cell wall, plasma membrane, chloroplast, membrane; EXPRESSED IN: 19 plant structures; EXPRESSED DURING: 10 growth stages; CONTAINS InterPro DOMAIN/s: Heat shock protein 70, conserved site (InterPro:IPR018181), Heat shock protein Hsp70 (InterPro:IPR001023), Heat shock protein 70 (InterPro:IPR013126); BEST Arabidopsis thaliana protein match is: HSP70 (heat shock protein 70); ATP binding (TAIR:AT3G12580.1); Has 24709 Blast hits to 24439 proteins in 3097 species: Archae - 107; Bacteria - 9686; Metazoa - 3084; Fungi - 1225; Plants - 724; Viruses - 242; Other Eukaryotes - 9641 (source: NCBI BLink). |
| GmaAffx.88762.1.S1_ | 0.02316 | 6.36092 | No Soybean Match Id | AT3G09270.1 | Encodes glutathione transferase belonging to the tau class of GSTs. Naming convention according to Wagner et al. (2002). |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.8386.1.S1_at | 0.02679 | 6.08033 | Glyma04g37040.1 | AT1G76650.1 | CALMODULIN-LIKE 38 (CML38); FUNCTIONS IN: calcium ion binding; INVOLVED IN: response to wounding; LOCATED IN: plasma membrane; EXPRESSED IN: 19 plant structures; EXPRESSED DURING: 12 growth stages; CONTAINS InterPro DOMAIN/s: EF-HAND 1 (InterPro:IPR018247), EF-HAND 2 (InterPro:IPR018249), Calcium-binding EF-hand (InterPro:IPR002048), EF-Hand type (InterPro:IPR011992), EF hand (InterPro:IPR018248); BEST Arabidopsis thaliana protein match is: calmodulin-related protein, putative (TAIR:AT1G76640.1); Has 11212 Blast hits to 8813 proteins in 1020 species: Archae - 0; Bacteria - 34; Metazoa - 5021; Fungi - 1984; Plants - 2417; Viruses - 0; Other Eukaryotes - 1756 (source: NCBI BLink). |
| GmaAffx.91561.1.S1_ | 0.05811 | 6.01771 | Glyma15g26590.1 | AT5G58070.1 | Encodes a temperature-induced lipocalin TIL1. Involved in thermotolerance. Peripherally associated with plasma membrane. |
| GmaAffx.6438.1.S1_a | 0.0678 | 5.92562 | Glyma11g29720.1 | AT2G38470.1 | Member of the plant WRKY transcription factor family. Regulates the antagonistic relationship between defense pathways mediating responses to P. syringae and necrotrophic fungal pathogens. Located in nucleus. Involved in response to various abiotic stresses - especially salt stress. |
| GmaAffx.92590.1.S1_ | 0.07663 | 5.84822 | Glyma02g33070.1 | AT2G41680.1 | Encodes a NADPH thioredoxin reductase involved in chloroplast protection against oxidative damage. |
| Gma.16015.1.S1_a_a | 0.08446 | 5.52025 | Glyma10g27610.1 | AT2G05620.1 | Involved in electron flow in Photosystem I. Essential for photoprotection. |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.13195.1.S1_s_a | 0.02743 | 5.44499 | Glyma01g20860.1 | AT1G13340.1 | | unknown protein; INVOLVED IN: response to oxidative stress; EXPRESSED IN: 11 plant structures; EXPRESSED DURING: 9 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF292, eukaryotic (InterPro:IPR005061); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT1G34220.2); Has 416 Blast hits to 416 proteins in 110 species: Archae - 0; Bacteria - 0; Metazoa - 148; Fungi - 88; Plants - 145; Viruses - 0; Other Eukaryotes - 35 (source: NCBI BLink). |
| GmaAffx.6519.1.S1_a | 0.06238 | 5.40828 | Glyma08g29470.1 | AT1G13340.1 | | unknown protein; INVOLVED IN: response to oxidative stress; EXPRESSED IN: 11 plant structures; EXPRESSED DURING: 9 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF292, eukaryotic (InterPro:IPR005061); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT1G34220.2); Has 416 Blast hits to 416 proteins in 110 species: Archae - 0; Bacteria - 0; Metazoa - 148; Fungi - 88; Plants - 145; Viruses - 0; Other Eukaryotes - 35 (source: NCBI BLink). |
| Gma.1622.1.A1_s_at | 0.02727 | 5.24619 | Glyma05g17470.1 | AT5G66900.1 | | disease resistance protein (CC-NBS-LRR class), putative; FUNCTIONS IN: protein binding, ATP binding; INVOLVED IN: defense response, apoptosis; CONTAINS InterPro DOMAIN/s: Disease resistance, plant (InterPro:IPR014011), NB-ARC (InterPro:IPR002182), Leucine-rich repeat (InterPro:IPR001611); BEST Arabidopsis thaliana protein match is: disease resistance protein (CC-NBS-LRR class), putative (TAIR:AT5G66910.1); Has 22201 Blast hits to 13213 proteins in 513 species: Archae - 34; Bacteria - 1477; Metazoa - 3674; Fungi - 1389; Plants - 15962; Viruses - 2; Other Eukaryotes - 914 (source: NCBI BLink). |
| GmaAffx.88762.1.S1_ | 0.02316 | 5.1926 | No Soybean Match Id | AT3G09270.1 | | Encodes glutathione transferase belonging to the tau class of GSTs. Naming convention according to Wagner et al. (2002). |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.5820.1.S1_s_at | 0.02679 | 4.85614 | Glyma06g19810.1 | AT1G19020.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: response to oxidative stress; LOCATED IN: cellular_component unknown; EXPRESSED IN: 19 plant structures; EXPRESSED DURING: 9 growth stages; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT3G48180.1); Has 58 Blast hits to 58 proteins in 9 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 58; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| Gma.16613.1.S1_s_a | 0.04002 | 4.66995 | Glyma10g40400.1 | AT5G67450.1 | | Encodes zinc-finger protein. mRNA levels are elevated in response to low temperature, cold temperatures and high salt. The protein is localized to the nucleus and acts as a transcriptional repressor. |
| Gma.4089.1.S1_s_at | 0.03956 | 4.51804 | Glyma12g01580.2 | AT4G25200.1 | | AtHSP23.6-mito mRNA, nuclear gene encoding mitochondrial |
| Gma.13829.1.A1_at | 0.05603 | 4.44067 | Glyma03g07560.1 | AT3G11170.1 | | Chloroplastic enzyme responsible for the synthesis of 16:3 and 18:3 fatty acids from galactolipids, sulpholipids and phosphatidylglycerol. Uses ferredoxin as electron donor. Gene expression is induced by wounding in shoot and root. The wound-response in shoot is independent of jasmonic acid mediated pathway whereas the root response is mediated by jasmonic acid. |
| GmaAffx.93206.1.S1_ | 0.06896 | 4.40263 | Glyma15g40190.1 | AT1G78380.1 | | Encodes a glutathione transferase that is a member of Tau GST gene family. Expression is induced by drought stress, oxidative stress, and high doses of auxin and cytokinin. naming convention according to Wagner et al. (2002) The expression of this gene is upregulated by herbicide safeners such as benoxacor and fenclorim. |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.10990.1.S1_at | 0.02679 | 4.38402 | Glyma04g01020.2 | AT2G21330.1 | | fructose-bisphosphate aldolase, putative; FUNCTIONS IN: fructose-bisphosphate aldolase activity, catalytic activity; INVOLVED IN: response to cadmium ion, pentose-phosphate shunt; LOCATED IN: in 8 components; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Aldolase-type TIM barrel (InterPro:IPR013785), Fructose-bisphosphate aldolase, class-I (InterPro:IPR000741); BEST Arabidopsis thaliana protein match is: fructose-bisphosphate aldolase, putative (TAIR:AT4G38970.1); Has 4194 Blast hits to 4191 proteins in 673 species: Archae - 0; Bacteria - 436; Metazoa - 1009; Fungi - 2; Plants - 339; Viruses - 0; Other Eukaryotes - 2408 (source: NCBI BLink). |
| Gma.16613.1.S1_at | 0.03476 | 4.29775 | Glyma10g40400.1 | AT5G67450.1 | | Encodes zinc-finger protein. mRNA levels are elevated in response to low temperature, cold temperatures and high salt. The protein is localized to the nucleus and acts as a transcriptional repressor. |
| Gma.1439.1.S1_at | 0.03169 | 4.21711 | Glyma04g14800.3 | AT3G22370.1 | | Encodes AOX1a, an isoform of alternative oxidase that is expressed in rosettes, flowers, and root. The alternative oxidase of plant mitochondria transfers electrons from the ubiquinone pool to oxygen without energy conservations. It is regulated through transcriptional control and by pyruvate. Plays a role in shoot acclimation to low temperature. Also is capable of ameliorating reactive oxygen species production when the cytochrome pathway is inhibited. AOX1a also functions as a marker for mitochondrial retrograde response. |
| Gma.2730.1.S1_at | 0.05759 | 4.16959 | Glyma06g20730.1 | AT5G62480.1 | | Encodes glutathione transferase belonging to the tau class of GSTs. Naming convention according to Wagner et al. (2002). |
| Gma.4027.1.S1_at | 0.05376 | 4.08927 | Glyma13g01150.1 | AT5G57560.1 | | Encodes a cell wall-modifying enzyme, rapidly upregulated in response to environmental stimuli |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.71308.2.A1 | 0.03369 | 4.00298 | Glyma04g05500.2 | AT2G26150.1 | | member of Heat Stress Transcription Factor (Hsf) family. Involved in response to misfolded protein accumulation in the cytosol. Regulated by alternative splicing and non-sense-mediated decay. |
| Gma.620.1.S1_at | 0.06613 | 3.93078 | Glyma15g40190.1 | AT1G78380.1 | | Encodes a glutathione transferase that is a member of Tau GST gene family. Expression is induced by drought stress, oxidative stress, and high doses of auxin and cytokinin. naming convention according to Wagner et al. (2002) The expression of this gene is upregulated by herbicide safeners such as benoxacor and fenclorim. |
| GmaAffx.90883.1.S1 | 0.06843 | 3.91553 | Glyma10g04960.2 | AT2G37570.1 | | encodes a protein that can complement the salt-sensitive phenotype of a calcineurin (CaN)-deficient yeast mutant. This gene occurs in a single-copy and is 75% identical to tobacco SLT1 gene. |
| Gma.3702.1.S1_at | 0.08156 | 3.91065 | Glyma13g42210.1 | AT3G54420.1 | | encodes an EP3 chitinase that is expressed during somatic embryogenesis in 'nursing' cells surrounding the embryos but not in embryos themselves. The gene is also expressed in mature pollen and growing pollen tubes until they enter the receptive synergid, but not in endosperm and integuments as in carrot. Post-embryonically, expression is found in hydathodes, stipules, root epidermis and emerging root hairs. |
| GmaAffx.74588.1.S1 | 0.0398 | 3.80582 | Glyma12g34210.1 | AT3G20600.1 | | Required for non-race specific resistance to bacterial and fungal pathogens.Mediates systemic acquired resistance (SAR) response. |
| Gma.4089.1.S1_at | 0.02761 | 3.72232 | No Soybean Match Id | AT4G25200.1 | | AtHSP23.6-mito mRNA, nuclear gene encoding mitochondrial |

| | | | | | |
|---|---|---|---|---|---|
| | | | | | Encodes a dehydrin protein whose expression is induced early on in response to dehydration stress. This gene's expression to cold occurs in two waves, with early induction occurring within 1 h and secondary induction occurring 5 h after the beginning of cold stress. Expression is also induced in response to ABA but not in response to 2,4-D, BA, and GA3. ERD14 protein is capable of binding Ca2+, especially when the protein is phosphorylated. |
| GmaAffx.89705.1.S1 | 0.05948 | 3.72217 | Glyma04g01130.2 | AT1G76180.1 | |
| Gma.1680.1.S1_x_at | 0.04484 | 3.70671 | No Soybean Match Id | AT3G05890.1 | RARE-COLD-INDUCIBLE 2B (RCI2B); INVOLVED IN: response to cold; LOCATED IN: integral to membrane; EXPRESSED IN: 19 plant structures; EXPRESSED DURING: 10 growth stages; CONTAINS InterPro DOMAIN/s: Uncharacterised protein family UPF0057 (InterPro:IPR000612); BEST Arabidopsis thaliana protein match is: RCI2A (RARE-COLD-INDUCIBLE 2A) (TAIR:AT3G05880.1); Has 905 Blast hits to 905 proteins in 327 species: Archae - 0; Bacteria - 464; Metazoa - 38; Fungi - 185; Plants - 185; Viruses - 0; Other Eukaryotes - 33 (source: NCBI BLink). |
| Gma.1680.1.S1_at | 0.03842 | 3.52926 | No Soybean Match Id | AT3G05890.1 | RARE-COLD-INDUCIBLE 2B (RCI2B); INVOLVED IN: response to cold; LOCATED IN: integral to membrane; EXPRESSED IN: 19 plant structures; EXPRESSED DURING: 10 growth stages; CONTAINS InterPro DOMAIN/s: Uncharacterised protein family UPF0057 (InterPro:IPR000612); BEST Arabidopsis thaliana protein match is: RCI2A (RARE-COLD-INDUCIBLE 2A) (TAIR:AT3G05880.1); Has 905 Blast hits to 905 proteins in 327 species: Archae - 0; Bacteria - 464; Metazoa - 38; Fungi - 185; Plants - 185; Viruses - 0; Other Eukaryotes - 33 (source: NCBI BLink). |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.36259.1.S1_ | 0.05811 | 3.49346 | Glyma05g33200.1 | | AT2G02120.1 | Predicted to encode a PR (pathogenesis-related) protein. Belongs to the plant defensin (PDF) family with the following members: At1g75830/PDF1.1, At5g44420/PDF1.2a, At2g26020/PDF1.2b, At5g44430/PDF1.2c, At2g26010/PDF1.3, At1g19610/PDF1.4, At1g55010/PDF1.5, At2g02120/PDF2.1, At2g02100/PDF2.2, At2g02130/PDF2.3, At1g61070/PDF2.4, At5g63660/PDF2.5, At2g02140/PDF2.6, At5g38330/PDF3.1 and At4g30070/PDF3.2. |
| Gma.8204.1.A1_at | 0.07712 | 3.46234 | Glyma04g14800.3 | | AT3G22370.1 | Encodes AOX1a, an isoform of alternative oxidase that is expressed in rosettes, flowers, and root. The alternative oxidase of plant mitochondria transfers electrons from the ubiquinone pool to oxygen without energy conservations. It is regulated through transcriptional control and by pyruvate. Plays a role in shoot acclimation to low temperature. Also is capable of ameliorating reactive oxygen species production when the cytochrome pathway is inhibited. AOX1a also functions as a marker for mitochondrial retrograde response. |
| Gma.8458.1.S1_at | 0.09221 | 3.26498 | Glyma08g20190.1 | | AT1G55020.1 | lipoxygenase, a defense gene conferring resistance Xanthomonas campestris |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | | | | AtbZIP60 consists of a bZIP DNA binding domain followed by a putative transmembrane domain. GFP fusions containing the first 260 amino acids (AtbZIP60deltaC) are nuclear-localized. AtbZIP60 is upregulated by the addition of tunicamycin (ER stress response inductor), DTT (inhibitor of disulfide bond formation) and azetin-2-carboxylate (proline analog perturbing protein structure). Upon ER stress the protein is proteolyzed and the soluble part is translocalized into the nucleus. AtbZIP60deltaC can activate the promoters of the ER chaperones BiP1, BiP2 and BiP3 and CNX1 and CNX2 via binding to the ER stress response element (ERSE) and the plant unfolded protein response element(P-UPRE). It can also activate its own transcription. |
| GmaAffx.3568.1.S1_a | 0.0196 | 3.20838 | Glyma02g19870.1 | AT1G42990.1 | | |
| Gma.6370.2.S1_a_at | 0.06877 | 3.1929 | Glyma13g32920.1 | AT2G23970.1 | defense-related protein, putative; INVOLVED IN: defense response; LOCATED IN: endomembrane system; EXPRESSED IN: leaf whorl, petal, sepal, flower, pollen tube; EXPRESSED DURING: 4 anthesis, petal differentiation and expansion stage; CONTAINS InterPro DOMAIN/s: Glutamine amidotransferase type 1 (InterPro:IPR017926); BEST Arabidopsis thaliana protein match is: glutamine amidotransferase class-I domain-containing protein (TAIR:AT4G30550.1); Has 4262 Blast hits to 4262 proteins in 788 species: Archae - 195; Bacteria - 1606; Metazoa - 9; Fungi - 124; Plants - 51; Viruses - 3; Other Eukaryotes - 2274 (source: NCBI BLink). | |
| GmaAffx.85068.1.S1_ | 0.09223 | 3.13772 | Glyma06g21900.2 | AT5G61150.2 | Encodes highly hydrophilic protein involved in positively regulating FLC expression. Mutants are early flowering and show a loss of FLC expression in the absence of cold. | |
| GmaAffx.92499.1.S1_ | 0.06133 | 3.123 | Glyma19g43460.1 | AT3G04720.1 | Encodes a protein similar to the antifungal chitin-binding protein hevein from rubber tree latex. mRNA levels increase in response to ethylene and turnip crinkle virus infection. | |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.80951.1.S1_ | 0.06821 | 3.04208 | Glyma16g29750.1 | AT5G52640.1 | | Encodes a cytosolic heat shock protein AtHSP90.1. AtHSP90.1 interacts with disease resistance signaling components SGT1b and RAR1 and is required for RPS2-mediated resistance. |
| GmaAffx.84566.1.S1_ | 0.05394 | 3.01127 | Glyma18g49360.1 | AT3G28910.1 | | transcription factor myb homologue |
| Gma.12393.1.S1_a_a | 0.05811 | 2.94675 | Glyma08g18800.1 | AT3G14440.1 | | Encodes 9-cis-epoxycarotenoid dioxygenase, a key enzyme in the biosynthesis of abscisic acid. Regulated in response to drought and salinity. Expressed in roots, flowers and seeds. Localized to the chloroplast stroma and thylakoid membrane. |
| Gma.2540.2.S1_s_at | 0.01829 | 2.92009 | Glyma17g24160.1 | AT1G54410.1 | | dehydrin family protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: response to water, response to stress; LOCATED IN: chloroplast; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Dehydrin (InterPro:IPR000167); Has 570 Blast hits to 469 proteins in 114 species: Archae - 0; Bacteria - 12; Metazoa - 159; Fungi - 66; Plants - 127; Viruses - 0; Other Eukaryotes - 206 (source: NCBI BLink). |
| Gma.7922.1.A1_a_at | 0.04307 | 2.88 | Glyma05g02210.1 | AT5G62520.1 | | Encodes a protein with similarity to RCD1 but without the WWE domain. The protein does have a PARP signature upstream of the C-terminal protein interaction domain. The PARP signature may bind NAD+ and attach the ADP-ribose-moiety from NAD+ to the target molecule. Its presence suggests a role for the protein in ADP ribosylation. Up-regulated by NaCl. SRO5 and P5CDH (an overlapping gene in the antisense orientation) generate 24-nt and 21-nt siRNAs, which together are components of a regulatory loop controlling reactive oxygen species (ROS) production and stress response. |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.82910.1.S1 | 0.06371 | 2.83792 | Glyma05g32670.2 | AT5G64030.1 | | dehydration-responsive protein-related; INVOLVED IN: biological_process unknown; LOCATED IN: Golgi apparatus; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF248, methyltransferase putative (InterPro:IPR004159); BEST Arabidopsis thaliana protein match is: dehydration-responsive protein-related (TAIR:AT1G29470.2); Has 70587 Blast hits to 32345 proteins in 1325 species: Archae - 209; Bacteria - 10119; Metazoa - 24829; Fungi - 6650; Plants - 2805; Viruses - 645; Other Eukaryotes - 25330 (source: NCBI BLink). |
| GmaAffx.92919.1.S1 | 0.08784 | 2.67416 | Glyma05g05290.1 | AT5G47120.1 | | Encodes BI-1, a homolog of mammalian Bax inhibitor 1. Functions as an attenuator of biotic and abiotic types of cell death. Bax-induced cell death can be downregulated by ectopically expressing AtBI in planta. |
| GmaAffx.93596.1.S1 | 0.04002 | 2.67185 | Glyma17g03950.2 | AT5G49520.1 | | Encodes WRKY48, a member of the WRKY Transcription Factor. WRKY48 is a stress- and pathogen-induced transcriptional activator that represses plant basal defense. |
| GmaAffx.34450.1.S1 | 0.07344 | 2.66914 | Glyma05g05290.1 | AT5G47120.1 | | Encodes BI-1, a homolog of mammalian Bax inhibitor 1. Functions as an attenuator of biotic and abiotic types of cell death. Bax-induced cell death can be downregulated by ectopically expressing AtBI in planta. |
| GmaAffx.19934.1.S1 | 0.04164 | 2.59328 | Glyma13g21490.2 | AT5G03720.1 | | Member of Heat Stress Transcription Factor (Hsf) family. Expression is regulated by DREB2A and in turn HSFA3 regulates the expression of hsps Hsp18.1-CI and Hsp26.5-MII35S. Involved in establishing thermotolerence. |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.1338.1.S1_a | 0.09622 | 2.55493 | Glyma02g35210.1 | AT3G11820.1 | Encodes a syntaxin localized at the plasma membrane (SYR1, Syntaxin Related Protein 1, also known as SYP121, PENETRATION1/PEN1). SYR1/PEN1 is a member of the SNARE superfamily proteins. SNARE proteins are involved in cell signaling, vesicle traffic, growth and development. SYR1/PEN1 functions in positioning anchoring of the KAT1 K+ channel protein at the plasma membrane. Transcription is upregulated by abscisic acid, suggesting a role in ABA signaling. Also functions in non-host resistance against barley powdery mildew, Blumeria graminis sp. hordei. SYR1/PEN1 is a nonessential component of the preinvasive resistance against Colletotrichum fungus. Required for mlo resistance. |
| Gma.2821.2.S1_at | 0.05376 | 2.50519 | Glyma11g02770.1 | AT4G11650.1 | L-lactate dehydrogenase, putative; FUNCTIONS IN: in 6 functions; INVOLVED IN: response to salt stress, response to abscisic acid stimulus; LOCATED IN: plasma membrane; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: L-lactate dehydrogenase (InterPro:IPR011304), NAD(P)-binding (InterPro:IPR016040), L-lactate/malate dehydrogenase (InterPro:IPR001557), Lactate/malate dehydrogenase (InterPro:IPR001236), L-lactate dehydrogenase, active site (InterPro:IPR018177), Lactate dehydrogenase/glycoside hydrolase, family 4, C-terminal (InterPro:IPR015955); BEST Arabidopsis thaliana protein match is: PMDH2 (peroxisomal NAD-malate dehydrogenase 2); malate dehydrogenase (TAIR:AT5G09660.2); Has 9812 Blast hits to 9805 proteins in 2184 species: Archae - 185; Bacteria - 5417; Metazoa - 1010; Fungi - 354; Plants - 275; Viruses - 0; Other Eukaryotes - 2571 (source: NCBI BLink). |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.93382.1.S1_ | 0.02727 | 2.4997 | Glyma04g01130.1 | AT1G76180.1 | | Encodes a dehydrin protein whose expression is induced early on in response to dehydration stress. This gene's expression to cold occurs in two waves, with early induction occurring within 1 h and secondary induction occurring 5 h after the beginning of cold stress. Expression is also induced in response to ABA but not in response to 2,4-D, BA, and GA3. ERD14 protein is capable of binding Ca2+, especially when the protein is phosphorylated. |
| Gma.13433.1.A1_at | 0.06337 | 2.492 | Glyma06g12480.1 | AT2G38310.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT5G05440.1); Has 225 Blast hits to 221 proteins in 31 species: Archae - 0; Bacteria - 2; Metazoa - 0; Fungi - 0; Plants - 222; Viruses - 0; Other Eukaryotes - 1 (source: NCBI BLink). |
| Gma.4589.1.S1_s_at | 0.04338 | 2.45797 | Glyma11g35800.1 | AT3G10985.1 | | A senescence-associated gene whose expression is induced in response to treatment with Nep1, a fungal protein that causes necrosis. |
| GmaAffx.90536.1.S1_ | 0.03947 | 2.42009 | Glyma13g20200.2 | AT5G03240.3 | | encodes ubiquitin that is attached to proteins destined for degradation. UBQ3 is most homologous with UBQ4, and is expressed in higher levels in vegetative tissue but lower levels in flowers than UBQ4. UBQ3 encodes different number of ubiquitins in different ecotypes. UBQ3 transcript level is modulated by UV-B and light/dark treatments. |
| Gma.8336.1.S1_at | 0.04669 | 2.39813 | Glyma19g40560.1 | AT2G47260.1 | | Encodes a member of WRKY Transcription Factor; Group I. Involved in nematode feeding site establishment. |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.93424.1.S1_ | 0.03697 | 2.38456 | Glyma10g05830.1 | AT5G03240.3 | | encodes ubiquitin that is attached to proteins destined for degradation. UBQ3 is most homologous with UBQ4, and is expressed in higher levels in vegetative tissue but lower levels in flowers than UBQ4. UBQ3 encodes different number of ubiquitins in different ecotypes. UBQ3 transcript level is modulated by UV-B and light/dark treatments. |
| Gma.4639.1.A1_at | 0.0963 | 2.28122 | Glyma13g36340.1 | AT3G20600.1 | | Required for non-race specific resistance to bacterial and fungal pathogens.Mediates systemic acquired resistance (SAR) response. |
| Gma.6796.1.S1_at | 0.06582 | 2.23799 | Glyma16g16290.1 | AT1G47128.1 | | cysteine proteinase precursor-like protein/ dehydration stress-responsive gene (RD21) |
| GmaAffx.91687.1.S1_ | 0.04783 | 2.21117 | Glyma08g39390.1 | AT1G60420.1 | | Reduce transmission through pollen. |
| GmaAffx.22821.1.S1_ | 0.0687 | 2.1784 | Glyma19g44390.2 | AT3G03300.2 | | Encodes a Dicer-like protein that functions in the antiviral silencing response in turnip-crinkle virus-infected plants but not in TMV or CMV-strain-Y-infected plants. Involved in the production of ta-siRNAs. Partially antagonizes the production of miRNAs by DCL1. Substitutes for DCL4 to produce viral siRNA when DCL4 is missing or inhibited. Able to produce siRNAs but not miRNAs. |
| GmaAffx.87406.1.S1_ | 0.06896 | 2.1552 | Glyma15g43040.1 | AT5G05170.1 | | Encodes a cellulose synthase isomer. CESA3 mutants have cellulose defect in the primary cell wall. Multiple lines of evidence suggest that CESA3, along with CESA1 and CESA6 are present in the same plasma membrane complex for cellulose biosynthesis. |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.11115.2.S1_at | 0.02679 | 2.14633 | Glyma11g14950.1 | | AT3G12580.1 | heat shock protein 70 (HSP70); FUNCTIONS IN: ATP binding; INVOLVED IN: in 8 processes; LOCATED IN: cytosol, mitochondrion, cell wall, plasma membrane; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 11 growth stages; CONTAINS InterPro DOMAIN/s: Heat shock protein 70, conserved site (InterPro:IPR018181), Heat shock protein Hsp70 (InterPro:IPR001023), Heat shock protein 70 (InterPro:IPR013126); BEST Arabidopsis thaliana protein match is: HSC70-1 (HEAT SHOCK COGNATE PROTEIN 70-1); ATP binding (TAIR:AT5G02500.1); Has 24865 Blast hits to 24554 proteins in 3102 species: Archae - 103; Bacteria - 9662; Metazoa - 3184; Fungi - 1196; Plants - 727; Viruses - 243; Other Eukaryotes - 9750 (source: NCBI BLink). |
| GmaAffx.11781.1.S1_ | 0.02727 | 2.13669 | Glyma02g35660.1 | | AT5G06320.1 | encodes a protein whose sequence is similar to tobacco hairpin-induced gene (HIN1) and Arabidopsis non-race specific disease resistance gene (NDR1). Expression of this gene is induced by cucumber mosaic virus, spermine and Pseudomonas syringae pv. tomato DC3000. The gene product is localized to the plasma membrane. |
| GmaAffx.85368.1.S1 | 0.06802 | 2.12413 | Glyma09g36540.1 | | AT5G61500.1 | ATG3; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: autophagy; LOCATED IN: cellular_component unknown; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Autophagocytosis associated protein, C-terminal (InterPro:IPR007135), Autophagocytosis associated protein, N-terminal (InterPro:IPR007134); Has 348 Blast hits to 341 proteins in 164 species: Archae - 0; Bacteria - 0; Metazoa - 118; Fungi - 114; Plants - 34; Viruses - 0; Other Eukaryotes - 82 (source: NCBI BLink). |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.3990.1.S1_at | 0.09531 | 2.12392 | Glyma05g31190.1 | AT2G22300.2 | Encodes a putative CAM binding transcription factor. Loss of function mutations show enhanced resistance to fungal and bacterial pathogens suggesting that CAMTA functions to suppress defense responses. |
| Gma.5260.1.A1_at | 0.05654 | 2.10446 | Glyma04g42740.1 | AT1G44170.2 | Encodes a protein similar to the aldehyde dehydrogenase cp-ADH from C.plantagineum. Constitutively expressed at low levels. |
| GmaAffx.89654.1.A1 | 0.03377 | 2.06334 | No Soybean Match Id | AT1G21750.1 | Encodes a protein disulfide isomerase-like (PDIL) protein, a member of a multigene family within the thioredoxin (TRX) superfamily; isoform contains non-consensus GA donor splice site at intron 9. Transcript levels for this gene are up-regulated in response to three different chemical inducers of ER stress (dithiothreitol, beta-mercaptoethanol, and tunicamycin). Neither AtIRE1-2 nor AtbZIP60 appear to be required for this response. |
| GmaAffx.2376.1.S1_a | 0.03623 | 2.04372 | Glyma04g06610.1 | AT2G25140.1 | Encodes ClpB4, which belongs to the Casein lytic proteinase/heat shock protein 100 (Clp/Hsp100) family. Targeted to the mitochondrion, also referred to as ClpB-m. Transcripts of ClpB4 accumulate dramatically at high temperatures, suggesting that it may be involved in response to heat stress. |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.20155.1.S1_ | 0.05171 | 2.03643 | Glyma02g35230.1 | AT3G11820.1 | | Encodes a syntaxin localized at the plasma membrane (SYR1, Syntaxin Related Protein 1, also known as SYP121, PENETRATION1/PEN1). SYR1/PEN1 is a member of the SNARE superfamily proteins. SNARE proteins are involved in cell signaling, vesicle traffic, growth and development. SYR1/PEN1 functions in positioning anchoring of the KAT1 K+ channel protein at the plasma membrane. Transcription is upregulated by abscisic acid, suggesting a role in ABA signaling. Also functions in non-host resistance against barley powdery mildew, Blumeria graminis sp. hordei. SYR1/PEN1 is a nonessential component of the preinvasive resistance against Colletotrichum fungus. Required for mlo resistance. |
| Gma.17785.1.S1_at | 0.0939 | 2.02414 | Glyma15g08370.1 | AT3G15210.1 | | Encodes a member of the ERF (ethylene response factor) subfamily B-1 of ERF/AP2 transcription factor family (ATERF-4). The protein contains one AP2 domain. Acts as a negative regulator of JA-responsive defense gene expression and resistance to the necrotrophic fungal pathogen Fusarium oxysporum and antagonizes JA inhibition of root elongation. |
| Gma.17631.1.S1_at | 0.06821 | 2.0009 | Glyma05g36620.1 | AT5G42020.1 | | Luminal binding protein (BiP2) involved in polar nuclei fusion during proliferation of endosperm nuclei. |
| Gma.16707.1.S1_a_a | 0.08379 | 1.99891 | Glyma03g37530.1 | AT3G05700.1 | | INVOLVED IN: response to water deprivation; CONTAINS InterPro DOMAIN/s: Drought induced 19 (InterPro:IPR008598); BEST Arabidopsis thaliana protein match is: drought-responsive family protein (TAIR:AT5G26990.1); Has 167 Blast hits to 167 proteins in 27 species: Archae - 0; Bacteria - 0; Metazoa - 38; Fungi - 0; Plants - 128; Viruses - 0; Other Eukaryotes - 1 (source: NCBI BLink). |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.15031.1.S1_ | 0.05218 | 1.97587 | Glyma12g04340.1 | AT4G39090.1 | | Similar to cysteine proteinases, induced by desiccation but not abscisic acid. Required for RRS1-R mediated resistance against Ralstonia solanacearum. Interacts with the R. solanacearum type III effector PopP2. RD19 associates with PopP2 to form a nuclear complex that is required for activation of the RRS1-R–mediated resistance response. |
| Gma.5914.1.S1_at | 0.06786 | 1.97422 | Glyma17g01400.1 | AT4G22120.5 | | early-responsive to dehydration protein-related / ERD protein-related; LOCATED IN: plasma membrane; EXPRESSED IN: 20 plant structures; EXPRESSED DURING: 11 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF221 (InterPro:IPR003864); BEST Arabidopsis thaliana protein match is: early-responsive to dehydration protein-related / ERD protein-related (TAIR:AT4G04340.3); Has 949 Blast hits to 842 proteins in 146 species: Archae - 0; Bacteria - 0; Metazoa - 161; Fungi - 455; Plants - 257; Viruses - 0; Other Eukaryotes - 76 (source: NCBI BLink). |
| GmaAffx.51116.1.S1_ | 0.0527 | 1.95335 | Glyma11g20840.1 | AT2G41540.3 | | Encodes a protein with NAD-dependent glycerol-3-phosphate (G3P) dehydrogenase which was shown to complement an Escherichia coli strain: BB20-14, auxotrophic for glycerol/G3P due to a loss-of-function mutation in the gpsA gene. |
| Gma.12045.1.S1_at | 0.09622 | 1.91789 | Glyma18g06840.1 | AT3G47340.1 | | encodes a glutamine-dependent asparagine synthetase, the predicted ASN1 peptide contains a purF-type glutamine-binding domain, and is expressed predominantly in shoot tissues, where light has a negative effect on its mRNA accumulation. Expression is induced within 3 hours of dark treatment, in senescing leaves and treatment with exogenous photosynthesis inhibitor. Induction of gene expression was suppressed in excised leaves supplied with sugar. The authors suggest that the gene's expression pattern is responding to the level of sugar in the cell. |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.80188.1.A1 | 0.08318 | 1.91332 | Glyma02g31490.1 | AT3G57530.1 | | Calcium-dependent Protein Kinase. ABA signaling component that regulates the ABA-responsive gene expression via ABF4. AtCPK32 has autophosphorylation activity and can phosphorylate ABF4 in vitro |
| Gma.2952.1.S1_at | 0.03169 | 1.89069 | Glyma14g26660.1 | AT1G77000.1 | | AtSKP2;2 is a homolog of human SKP2, the human F-box protein that recruits E2F1. Contains an F-box motif at the N-terminal region and a C-terminal Leu-rich repeat domain. Forms part of an E3-ubiquitin-ligase SCF (Skp1, cullin, F-box) complex and recruits phosphorylated AtE2Fc, a transcriptional factor that might play a role in cell division and during the transition from skotomorphogenesis to photomorphogenesis. AtSKP2;1 (At1g21410) and AtSKP2;2 (At1g77000) may be duplicated genes. AtSKP2b may also be involved in the degradation of KRP1/ICK1, a CDK inhibitor. |
| Gma.54.1.S1_at | 0.04618 | 1.86771 | Glyma01g41260.1 | AT4G33950.1 | | Encodes calcium-independent ABA-activated protein kinase, a member of SNF1-related protein kinases (SnRK2) whose activity is activated by ionic (salt) and non-ionic (mannitol) osmotic stress. Mutations disrupted ABA induction of stomatal closure as well as ABA inhibition of light-induced stomatal opening. However, regulation of stomatal opening/closing by light or CO(2) is not affected in these mutants. May act in the interval between ABA perception and reactive oxygen species production in the ABA signalling network. |
| Gma.1112.1.S1_s_at | 0.03172 | 1.84656 | Glyma09g38990.2 | AT1G01470.1 | | Encodes late-embryogenesis abundant protein whose mRNA levels are induced in response to wounding and light stress. Might be involved in protection against dessication. |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.4319.1.S1_s_at | 0.09624 | 1.84268 | No Soybean Match Id | AT3G07230.1 | | wound-responsive protein-related; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Wound-inducible basic (InterPro:IPR012643); Has 20 Blast hits to 20 proteins in 8 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 20; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| Gma.6803.2.S1_s_at | 0.06238 | 1.82426 | Glyma12g05150.4 | AT3G50830.1 | | cold acclimation protein WCOR413-like protein beta form. Transcript is not detectable. |
| GmaAffx.69792.1.S1_ | 0.08446 | 1.80964 | Glyma16g27950.1 | AT5G64750.1 | | Encodes a putative transcription factor containing an AP2 domain. Is a member of the ERF (ethylene response factor) subfamily B-4 of ERF/AP2 transcription factor family. Expressed in response to ABA, osmotic stress, sugar stress and drought. Mutants are hypersensitive to these stresses. May be involved in regulation of ABA mediated stress response. |
| Gma.10639.1.S1_x_a | 0.06499 | 1.79282 | Glyma03g27560.1 | AT2G41010.1 | | Encodes a novel calmodulin binding protein whose gene expression is induced by dehydration and ionic (salt) and non-ionic (mannitol) osmotic stress. Lines over-expressing this gene are more sensitive and anti-sense lines are more tolerant to osmotic stress, suggesting this gene may be a negative regulator of response to osmotic stress. |
| Gma.6474.1.A1_s_at | 0.06552 | 1.72727 | Glyma15g13470.1 | AT2G34690.1 | | Gene product transports the glycolipid precursor sphingosine between membranes in vitro. Mutant constitutively expresses defense-related genes that accompany the hypersensitive response normally triggered by avirulent pathogens. |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.10850.2.S1_at | 0.02727 | 1.71445 | Glyma04g01130.1 | AT1G76180.1 | | Encodes a dehydrin protein whose expression is induced early on in response to dehydration stress. This gene's expression to cold occurs in two waves, with early induction occurring within 1 h and secondary induction occurring 5 h after the beginning of cold stress. Expression is also induced in response to ABA but not in response to 2,4-D, BA, and GA3. ERD14 protein is capable of binding Ca2+, especially when the protein is phosphorylated. |
| GmaAffx.1991.1.S1_a | 0.06366 | 1.66499 | Glyma11g04040.2 | AT5G47120.1 | | Encodes Bl-1, a homolog of mammalian Bax inhibitor 1. Functions as an attenuator of biotic and abiotic types of cell death. Bax-induced cell death can be downregulated by ectopically expressing AtBl in planta. |
| GmaAffx.53527.1.S1_ | 0.04484 | 1.64213 | Glyma11g09930.1 | AT3G21070.2 | | Encodes a protein with NAD(H) kinase activity. |
| Gma.6325.1.S1_at | 0.06238 | 1.63445 | Glyma17g24650.1 | AT1G29150.1 | | specifically interacts with FUS6/COP11 via the C-terminal domain of FUS6/COP11 and associates with an ATPase subunit of the 19S proteasome regulatory complex, AtS6A. |
| Gma.6803.1.S1_a_at | 0.03261 | 1.61499 | Glyma12g05150.4 | AT3G50830.1 | | cold acclimation protein WCOR413-like protein beta form. Transcript is not detectable. |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.5594.1.S1_at | 0.09471 | 1.60659 | Glyma03g33990.1 | AT3G53230.1 | cell division cycle protein 48, putative / CDC48, putative; FUNCTIONS IN: hydrolase activity, nucleoside-triphosphatase activity, binding, nucleotide binding, ATP binding; INVOLVED IN: response to cadmium ion; LOCATED IN: cytosol, nucleolus, plasma membrane; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: ATPase, AAA+ type, core (InterPro:IPR003593), ATPase, AAA-type, core (InterPro:IPR003959), ATPase, AAA-type, conserved site (InterPro:IPR003960), Aspartate decarboxylase-like fold (InterPro:IPR009010), Cell division protein 48, CDC48, domain 2 (InterPro:IPR004201), ATPase, AAA-type, VAT, N-terminal (InterPro:IPR003338), ATPase, AAA-type, CDC48 (InterPro:IPR005938); BEST Arabidopsis thaliana protein match is: cell division cycle protein 48, putative / CDC48, putative (TAIR:AT5G03340.1); Has 44293 Blast hits to 25363 proteins in 1885 species: Archae - 996; Bacteria - 15540; Metazoa - 7206; Fungi - 4097; Plants - 3052; Viruses - 60; Other Eukaryotes - 13342 (source: NCBI BLink). |
| GmaAffx.65393.1.S1_ | 0.09458 | 1.58233 | Glyma10g02210.1 | AT4G02380.1 | Encodes AtLEA5 (late embryogenesis abundant like protein). Also known as SENESCENCE-ASSOCIATED GENE 21 (SAG21). Has a role on oxidative stress tolerance. mRNA levels are elevated in response to various stresses. |
| GmaAffx.80733.1.S1_ | 0.03964 | 1.57789 | Glyma17g11240.1 | AT5G20320.1 | Encodes an RNase III-like enzyme that catalyzes processing of trans-acting small interfering RNA precursors in a distinct small RNA biogenesis pathway. The protein is also involved in the production of 21-nt primary siRNAs from both inverted-repeat constructs and endogenous sequences, as well as the RDR6-dependent 21-nt secondary siRNAs involved in long-range cell-to-cell signaling. It binds DRB4, a ds-RNA binding protein. |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.1452.1.S1_at | 0.0863 | 1.53216 | Glyma06g04460.1 | | AT1G49760.1 | polyadenylate-binding protein, putative / PABP, putative, similar to poly(A)-binding protein GB:AAF66825 GI:7673359 from (Nicotiana tabacum). Highly and ubiquitously expressed. Member of the class II PABP family. |
| GmaAffx.93598.1.S1_ | 0.08784 | 1.52902 | Glyma09g38990.2 | | AT1G01470.1 | Encodes late-embryogenesis abundant protein whose mRNA levels are induced in response to wounding and light stress. Might be involved in protection against dessication. |
| Gma.1917.1.S1_s_at | 0.09268 | 1.51813 | Glyma08g18690.2 | | AT1G78380.1 | Encodes a glutathione transferase that is a member of Tau GST gene family. Expression is induced by drought stress, oxidative stress, and high doses of auxin and cytokinin. naming convention according to Wagner et al. (2002) The expression of this gene is upregulated by herbicide safeners such as benoxacor and fenclorim. |
| Gma.4192.1.S1_at | 0.09671 | 1.48986 | Glyma05g01970.1 | | AT3G48030.1 | hypoxia-responsive family protein / zinc finger (C3HC4-type RING finger) family protein; FUNCTIONS IN: protein binding, zinc ion binding; INVOLVED IN: response to hypoxia; EXPRESSED IN: hypocotyl, root; CONTAINS InterPro DOMAIN/s: Zinc finger, RING-type (InterPro:IPR001841), Zinc finger, RING/FYVE/PHD-type (InterPro:IPR013083), Hypoxia induced protein conserved region (InterPro:IPR007667); BEST Arabidopsis thaliana protein match is: zinc finger (C3HC4-type RING finger) family protein (TAIR:AT1G23980.1); Has 6706 Blast hits to 6689 proteins in 224 species: Archae - 0; Bacteria - 6; Metazoa - 2119; Fungi - 610; Plants - 2790; Viruses - 41; Other Eukaryotes - 1140 (source: NCBI BLink). |
| GmaAffx.52838.1.S1_ | 0.03169 | 1.48759 | Glyma02g10360.1 | | AT2G42010.1 | phospholipase D (PLDbeta) |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.5553.2.S1_at | 0.06304 | 1.4827 | Glyma01g26300.1 | AT3G53990.1 | | universal stress protein (USP) family protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: response to cold, response to stress; LOCATED IN: cellular_component unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: UspA (InterPro:IPR006016), Rossmann-like alpha/beta/alpha sandwich fold (InterPro:IPR014729), Universal stress protein A (InterPro:IPR006015); BEST Arabidopsis thaliana protein match is: universal stress protein (USP) family protein / early nodulin ENOD18 family protein (TAIR:AT3G03270.2); Has 2509 Blast hits to 2475 proteins in 563 species: Archae - 245; Bacteria - 1658; Metazoa - 47; Fungi - 64; Plants - 395; Viruses - 0; Other Eukaryotes - 100 (source: NCBI BLink). |
| Gma.3755.1.S1_at | 0.08694 | 1.48094 | Glyma03g30590.1 | AT5G13320.1 | | Encodes an enzyme capable of conjugating amino acids to 4-substituted benzoates. 4-HBA (4-hydroxybenzoic acid) and pABA (4-aminobenzoate) may be targets of the enzyme in Arabidopsis, leading to the production of pABA-Glu, 4HBA-Glu, or other related compounds. This enzyme is involved in disease-resistance signaling. It is required for the accumulation of salicylic acid, activation of defense responses, and resistance to Pseudomonas syringae. Salicylic acid can decrease this enzyme's activity in vitro and may act as a competitive inhibitor. Expression of PBS3/GH3.12 can be detected in cotyledons, true leaves, hypocotyls, and occasionally in some parts of roots from 10-day-old seedlings. No expression has been detected in root, stem, rosette or cauline leaves of mature 4- to 5-week-old plants. |
| GmaAffx.26356.1.S1_ | 0.05811 | 1.47572 | Glyma09g11770.2 | AT1G30270.2 | | |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.69748.1.S1_ | 0.07532 | 1.47391 | No Soybean Match Id | AT1G02270.1 | endonuclease/exonuclease/phosphatase family protein / calcium binding EF hand family protein; FUNCTIONS IN: catalytic activity, calcium ion binding; INVOLVED IN: response to cold; LOCATED IN: nucleus; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: EF-HAND 1 (InterPro:IPR018247), EF-HAND 2 (InterPro:IPR018249), EF-Hand type (InterPro:IPR011992), EF hand (InterPro:IPR018248), Endonuclease/exonuclease/phosphatase (InterPro:IPR005135); BEST Arabidopsis thaliana protein match is: calcium-binding EF hand family protein (TAIR:AT5G54130.2); Has 335 Blast hits to 335 proteins in 87 species: Archae - 0; Bacteria - 0; Metazoa - 134; Fungi - 41; Plants - 110; Viruses - 0; Other Eukaryotes - 50 (source: NCBI BLink). |
| Gma.10881.1.S1_at | 0.0819 | 1.43541 | Glyma06g00990.1 | AT4G34710.2 | encodes a arginine decarboxylase (ADC), a rate-limiting enzyme that catalyzes the first step of polyamine (PA) biosynthesis via ADC pathway in Arabidopsis thaliana. Arabidopsis genome has two ADC paralogs, ADC1 and ADC2. ADC2 is stress-inducible (osmotic stress). Double mutant analysis showed that ADC genes are essential for the production of PA, and are required for normal seed development. Overexpression causes phenotypes similar to GA-deficient plants and these plants show reduced levels of GA due to lower expression levels of AtGA20ox1, AtGA3ox3 and AtGA3ox1. |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.19777.1.A1 | 0.09858 | 1.42994 | No Soybean Match Id | AT4G31800.2 | | Pathogen-induced transcription factor. Binds W-box sequences in vitro. Forms protein complexes with itself and with WRKY40 and WRKY60. Constitutive expression of WRKY18 enhanced resistance to P. syringae, but its coexpression with WRKY40 or WRKY60 made plants more susceptible to both P. syringae and B. cinerea. WRKY18, WRKY40, and WRKY60 have partially redundant roles in response to the hemibiotrophic bacterial pathogen Pseudomonas syringae and the necrotrophic fungal pathogen Botrytis cinerea, with WRKY18 playing a more important role than the other two. |
| GmaAffx.48022.2.A1 | 0.06582 | 1.42933 | Glyma13g02470.3 | AT4G08500.1 | | Encodes a member of the A1 subgroup of the MEKK (MAPK/ERK kinase kinase) family. MEKK is another name for Mitogen-Activated Protein Kinase Kinase (MAPKKK or MAP3K). This subgroup has four members: At4g08500 (MEKK1, also known as ARAKIN, MAP3Kb1, MAPKKK8), At4g08480 (MEKK2, also known as MAP3Kb4, MAPKKK9), At4g08470 (MEKK3, also known as MAP3Kb3, MAPKKK10) and At4g12020 (MEKK4, also known as MAP3Kb5, MAPKKK11, WRKY19). Nomenclatures for mitogen-activated protein kinases are described in Trends in Plant Science 2002, 7(7):301. Mediates cold, salt, cadmium and wounding stress signalling. Phosphorylates MEK1. |
| Gma.17692.1.S1_at | 0.047 | 1.41779 | Glyma06g24500.1 | AT1G52740.1 | | Encodes HTA9, a histone H2A protein. |
| Gma.10796.2.S1_at | 0.0958 | 1.3859 | Glyma09g15620.1 | AT5G05170.1 | | Encodes a cellulose synthase isomer. CESA3 mutants have cellulose defect in the primary cell wall. Multiple lines of evidence suggest that CESA3, along with CESA1 and CESA6 are present in the same plasma membrane complex for cellulose biosynthesis. |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.18441.1.S1_ | 0.0982 | 1.37488 | Glyma13g28090.1 | AT3G23600.1 | | dienelactone hydrolase family protein; FUNCTIONS IN: hydrolase activity; INVOLVED IN: response to salt stress; LOCATED IN: apoplast, nucleus, plasma membrane, cytoplasm; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Dienelactone hydrolase (InterPro:IPR002925); BEST Arabidopsis thaliana protein match is: dienelactone hydrolase family protein (TAIR:AT3G23570.1); Has 675 Blast hits to 673 proteins in 155 species: Archae - 2; Bacteria - 132; Metazoa - 54; Fungi - 330; Plants - 132; Viruses - 0; Other Eukaryotes - 25 (source: NCBI BLink). |
| GmaAffx.1917.1.S1_s | 0.06238 | 1.36688 | Glyma08g41220.2 | AT1G26850.2 | | dehydration-responsive family protein; LOCATED IN: Golgi apparatus, membrane; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF248, methyltransferase putative (InterPro:IPR004159); BEST Arabidopsis thaliana protein match is: dehydration-responsive family protein (TAIR:AT4G18030.1); Has 516 Blast hits to 509 proteins in 51 species: Archae - 0; Bacteria - 55; Metazoa - 0; Fungi - 0; Plants - 457; Viruses - 0; Other Eukaryotes - 4 (source: NCBI BLink). |
| Gma.10796.2.S1_x_a | 0.06802 | 1.3531 | Glyma09g15620.1 | AT5G05170.1 | | Encodes a cellulose synthase isomer. CESA3 mutants have cellulose defect in the primary cell wall. Multiple lines of evidence suggest that CESA3, along with CESA1 and CESA6 are present in the same plasma membrane complex for cellulose biosynthesis. |
| Gma.10796.1.S1_a_a | 0.03375 | 1.34315 | Glyma15g43040.1 | AT5G05170.1 | | Encodes a cellulose synthase isomer. CESA3 mutants have cellulose defect in the primary cell wall. Multiple lines of evidence suggest that CESA3, along with CESA1 and CESA6 are present in the same plasma membrane complex for cellulose biosynthesis. |

| | | | | | |
|---|---|---|---|---|---|
| Gma.544.1.A1_at | 0.07148 | 1.30298 | Glyma05g05580.1 | AT5G47100.1 | member of AtCBLs (Calcineurin B-like Calcium Sensor Proteins. CBL9 interacts with and targets CIPK23 to the plasma membrane in vivo. |
| Gma.13118.1.S1_at | 0.04862 | 1.2755 | Glyma05g25320.4 | AT3G48750.1 | A-type cyclin-dependent kinase. Together with its specific inhibitor, the Kip-related protein, KRP2 they regulate the mitosis-to-endocycle transition during leaf development. Dominant negative mutations abolish cell division. Loss of function phenotype has reduced fertility with failure to transmit via pollen. Pollen development is arrested at the second mitotic division. Expression is regulated by environmental and chemical signals. Part of the promoter is responsible for expression in trichomes. Functions as a positive regulator of cell proliferation during development of the male gametophyte, embryo and endosperm. Phosphorylation of threonine 161 is required for activation of its associated kinase. |
| Gma.2741.1.S1_at | 0.09193 | 1.24966 | Glyma12g09680.1 | AT5G22360.1 | Member of Synaptobrevin-like AtVAMP7C, v-SNARE protein family. |
| Gma.18032.1.S1_at | 0.07344 | 1.11459 | Glyma14g40320.4 | AT4G24190.2 | encodes an ortholog of GRP94, an ER-resident HSP90-like protein and is involved in regulation of meristem size and organization. Single and double mutant analyses suggest that SHD may be required for the correct folding and/or complex formation of CLV proteins. Lines carrying recessive mutations in this locus exhibits expanded shoot meristems, disorganized root meristems, and defective pollen tube elongation. Transcript is detected in all tissues examined and is not induced by heat. Endoplasmin supports the protein secretory pathway and has a role in proliferating tissues. |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.5283.1.S1_at | 0.03707 | -17.624 | Glyma06g07300.1 | AT2G18660.1 | | Encodes PNP-A (Plant Natriuretic Peptide A). PNPs are a class of systemically mobile molecules distantly related to expansins; their biological role has remained elusive. PNP-A contains a signal peptide domain and is secreted into the extracellular space. Co-expression analyses using microarray data suggest that PNP-A may function as a component of plant defence response and SAR in particular, and could be classified as a newly identified PR protein. |
| GmaAffx.39920.1.S1_ | 0.07766 | -9.6103 | Glyma15g11950.1 | AT4G22260.1 | | Similar to mitochondrial alternative oxidase. im mutants have a variegated phenotype and fail to differentiate chloroplasts in the majority of their cells under high light intensity continuous illumination. The white tissues of immutans accumulate phytoene, a non-colored C40 carotenoid intermediate. This suggests that immutans controls, either directly or indirectly, the activity of phytoene desaturase (PDS), the enzyme that converts phytoene to zeta-carotene in higher plants. However, im is not the structural gene for PDS. It is located in the lumenar face of the thylakoid membrane. IM is expressed ubiquitously in plant tissues. |
| Gma.5629.2.S1_a_at | 0.08118 | -5.4606 | Glyma15g05820.1 | AT2G41480.1 | | electron carrier/ heme binding / peroxidase; FUNCTIONS IN: electron carrier activity, peroxidase activity, heme binding; INVOLVED IN: response to oxidative stress; EXPRESSED IN: root, leaf; EXPRESSED DURING: LP.04 four leaves visible; CONTAINS InterPro DOMAIN/s: Haem peroxidase (InterPro:IPR010255), Plant peroxidase (InterPro:IPR000823), Haem peroxidase, plant/fungal/bacterial (InterPro:IPR002016); BEST Arabidopsis thaliana protein match is: peroxidase, putative (TAIR:AT5G64120.1); Has 2602 Blast hits to 2587 proteins in 148 species: Archae - 0; Bacteria - 4; Metazoa - 2; Fungi - 18; Plants - 2558; Viruses - 0; Other Eukaryotes - 20 (source: NCBI BLink). |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.84317.1.S1_ | 0.05376 | -5.4574 | Glyma06g42310.1 | AT5G54250.2 | | member of Cyclic nucleotide gated channel family, downstream component of the signaling pathways leading to HR resistance. mutant plants exhibit gene-for-gene disease resistance against avirulent Pseudomonas syringae despite the near-complete absence of the hypersensitive response (HR). Salicylic acid accumulation in dnd2 mutants is completely PAD4-independent. |
| Gma.5629.1.S1_at | 0.0317 | -5.0209 | No Soybean Match Id | AT2G41480.1 | | electron carrier/ heme binding / peroxidase; FUNCTIONS IN: electron carrier activity, peroxidase activity, heme binding; INVOLVED IN: response to oxidative stress; EXPRESSED IN: root, leaf; EXPRESSED DURING: LP.04 four leaves visible; CONTAINS InterPro DOMAIN/s: Haem peroxidase (InterPro:IPR010255), Plant peroxidase (InterPro:IPR000823), Haem peroxidase, plant/fungal/bacterial (InterPro:IPR002016); BEST Arabidopsis thaliana protein match is: peroxidase, putative (TAIR:AT5G64120.1); Has 2602 Blast hits to 2587 proteins in 148 species: Archae - 0; Bacteria - 4; Metazoa - 2; Fungi - 18; Plants - 2558; Viruses - 0; Other Eukaryotes - 20 (source: NCBI BLink). |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.5971.1.S1_at | 0.047 | -4.8126 | Glyma11g07670.1 | AT5G66390.1 | | peroxidase 72 (PER72) (P72) (PRXR8); FUNCTIONS IN: electron carrier activity, peroxidase activity, heme binding; INVOLVED IN: response to oxidative stress; LOCATED IN: endomembrane system; EXPRESSED IN: leaf apex, hypocotyl, flower, root; EXPRESSED DURING: 4 anthesis, petal differentiation and expansion stage; CONTAINS InterPro DOMAIN/s: Haem peroxidase (InterPro:IPR010255), Plant peroxidase (InterPro:IPR000823), Haem peroxidase, plant/fungal/bacterial (InterPro:IPR002016); BEST Arabidopsis thaliana protein match is: peroxidase, putative (TAIR:AT4G36430.1); Has 2940 Blast hits to 2923 proteins in 224 species: Archae - 0; Bacteria - 4; Metazoa - 0; Fungi - 117; Plants - 2778; Viruses - 0; Other Eukaryotes - 41 (source: NCBI BLink). |
| Gma.1778.1.S1_at | 0.05962 | -3.9747 | Glyma08g16220.1 | AT3G12610.1 | | Plays role in DNA-damage repair/toleration. Partially complements RecA- phenotypes. |
| GmaAffx.74124.1.S1 | 0.08292 | -3.9375 | No Soybean Match Id | AT5G67400.1 | | peroxidase 73 (PER73) (P73) (PRXR11); FUNCTIONS IN: electron carrier activity, peroxidase activity, heme binding; INVOLVED IN: response to oxidative stress; LOCATED IN: endomembrane system; EXPRESSED IN: root; CONTAINS InterPro DOMAIN/s: Haem peroxidase (InterPro:IPR010255), Plant peroxidase (InterPro:IPR000823), Haem peroxidase, plant/fungal/bacterial (InterPro:IPR002016); BEST Arabidopsis thaliana protein match is: peroxidase, putative (TAIR:AT3G49960.1); Has 2886 Blast hits to 2873 proteins in 211 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 83; Plants - 2764; Viruses - 0; Other Eukaryotes - 39 (source: NCBI BLink). |
| Gma.3562.1.S1_at | 0.03169 | -3.8892 | Glyma20g38580.1 | AT3G55120.1 | | Catalyzes the conversion of chalcones into flavanones. Required for the accumulation of purple anthocyanins in leaves and stems. |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.1539.1.S1_at | 0.06449 | -3.6772 | Glyma11g05300.2 | AT4G37520.1 | | peroxidase 50 (PER50) (P50) (PRXR2); FUNCTIONS IN: electron carrier activity, peroxidase activity, heme binding; INVOLVED IN: response to oxidative stress; LOCATED IN: cytoplasm; CONTAINS InterPro DOMAIN/s: Haem peroxidase (InterPro:IPR010255), Plant peroxidase (InterPro:IPR000823), Haem peroxidase, plant/fungal/bacterial (InterPro:IPR002016); BEST Arabidopsis thaliana protein match is: peroxidase, putative (TAIR:AT4G37530.1); Has 3003 Blast hits to 2988 proteins in 241 species: Archae - 0; Bacteria - 0; Metazoa - 2; Fungi - 172; Plants - 2780; Viruses - 0; Other Eukaryotes - 49 (source: NCBI BLink). |
| GmaAffx.73002.1.S1_ | 0.06017 | -3.5697 | Glyma02g09470.1 | AT1G14790.1 | | Encodes RNA-dependent RNA polymerase. While not required for virus-induced post-transcriptional gene silencing (PTGS), it can promote turnover of viral RNAs in infected plants. Nomenclature according to Xie, et al. (2004). Involved in the production of Cucumber Mosaic Virus siRNAs. |
| Gma.4126.1.S1_at | 0.05811 | -3.4612 | Glyma06g16810.1 | AT5G63660.1 | | Predicted to encode a PR (pathogenesis-related) protein. Belongs to the plant defensin (PDF) family with the following members: At1g75830/PDF1.1, At5g44420/PDF1.2a, At2g26020/PDF1.2b, At5g44430/PDF1.2c, At2g26010/PDF1.3, At1g19610/PDF1.4, At1g55010/PDF1.5, At2g02120/PDF2.1, At2g02100/PDF2.2, At2g02130/PDF2.3, At1g61070/PDF2.4, At5g63660/PDF2.5, At2g02140/PDF2.6, At5g38330/PDF3.1 and At4g30070/PDF3.2. |
| Gma.3881.2.S1_x_at | 0.04515 | -3.4266 | Glyma14g38580.1 | AT2G30490.1 | | Encodes a cinnamate-4-hydroxylase. Mutations in this gene impact phenylpropanoid metabolism, growth and development. |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.2436.1.A1_ | 0.02472 | -3.3928 | Glyma10g38590.1 | AT5G51820.1 | | Encodes a plastid isoform of the enzyme phosphoglucomutase involved in controlling photosynthetic carbon flow. Effective petiole movement against the direction of the gravity requires functional PGM activity that is required for full development of amyloplasts. |
| Gma.1445.1.S1_at | 0.04208 | -3.217 | Glyma13g44700.1 | AT1G15950.1 | | Encodes a cinnamoyl CoA reductase. Involved in lignin biosynthesis. |
| Gma.4919.1.S1_at | 0.02679 | -3.1978 | Glyma14g05840.1 | AT5G05340.1 | | peroxidase, putative; FUNCTIONS IN: electron carrier activity, peroxidase activity, heme binding; INVOLVED IN: response to oxidative stress; LOCATED IN: apoplast, cell wall; EXPRESSED IN: 9 plant structures; EXPRESSED DURING: 4 anthesis, LP.04 four leaves visible, C globular stage, petal differentiation and expansion stage; CONTAINS InterPro DOMAIN/s: Haem peroxidase (InterPro:IPR010255), Plant peroxidase (InterPro:IPR000823), Haem peroxidase, plant/fungal/bacterial (InterPro:IPR002016); BEST Arabidopsis thaliana protein match is: peroxidase, putative (TAIR:AT5G58400.1); Has 2945 Blast hits to 2930 proteins in 223 species: Archae - 0; Bacteria - 4; Metazoa - 2; Fungi - 110; Plants - 2778; Viruses - 0; Other Eukaryotes - 51 (source: NCBI BLink). |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.13045.1.S1_at | 0.08512 | -3.1291 | Glyma05g33630.1 | AT3G17790.1 | | PAP17; FUNCTIONS IN: phosphatase activity, protein serine/threonine phosphatase activity, acid phosphatase activity; INVOLVED IN: response to hydrogen peroxide, cellular phosphate ion homeostasis; LOCATED IN: cell surface; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Metallophosphoesterase (InterPro:IPR004843); BEST Arabidopsis thaliana protein match is: PAP3 (PURPLE ACID PHOSPHATASE 3); acid phosphatase/ protein serine/threonine phosphatase (TAIR:AT1G14700.1); Has 911 Blast hits to 906 proteins in 224 species: Archae - 2; Bacteria - 200; Metazoa - 324; Fungi - 6; Plants - 95; Viruses - 0; Other Eukaryotes - 284 (source: NCBI BLink). |
| Gma.17073.1.S1_at | 0.07385 | -3.0137 | Glyma01g24530.1 | AT2G39800.1 | | encodes a delta1-pyrroline-5-carboxylate synthase that catalyzes the rate-limiting enzyme in the biosynthesis of proline. Gene is expressed in reproductive organs and tissues under non-stress conditions but in the whole plant under water-limiting condition. Expression is also induced by abscisic acid and salt stress in a light-dependent manner. P5CS1 appears to be involved in salt stress responses related to proline accumulation, including protection from reactive oxidative species. P5CS1 appears to be present in different cells and/or different subcellular locations from P5CS2 in a tissue-dependent manner. |
| GmaAffx.88815.1.A1 | 0.06865 | -2.7354 | Glyma16g04240.3 | AT5G17920.2 | | Encodes a cytosolic cobalamin-independent methionine synthase, involved in methionine regeneration via the activated methyl cycle (SAM cycle). The protein undergoes thiolation following treatment with the oxidant tert-butylhydroperoxide. |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.21668.1.A1 | 0.04862 | -2.7304 | Glyma18g50750.1 | AT5G60920.1 | Encodes a glycosylphosphatidylinositol-anchored protein localized primarily in the plasma membrane of the longitudinal sides of root cells. Necessary for oriented cell expansion in Arabidopsis. Cob mutants have abnormal roots that expand radially rather than longitudinally under certain growth conditions. |
| Gma.1445.1.S1_a_at | 0.04466 | -2.7301 | No Soybean Match Id | AT1G80820.1 | Encodes an cinnamoyl CoA reductase isoform. Involved in lignin biosynthesis. |
| Gma.4928.1.S1_a_at | 0.03369 | -2.7133 | Glyma15g00390.1 | AT1G51680.1 | encodes an isoform of 4-coumarate:CoA ligase (4CL), which is involved in the last step of the general phenylpropanoid pathway. In addition to 4-coumarate, it also converts ferulate. The catalytic efficiency was in the following (descending) order: p-coumaric acid, ferulic acid, caffeic acid and 5-OH-ferulic acid. At4CL1 was unable to use sinapic acid as substrate. |
| Gma.3881.2.S1_at | 0.08118 | -2.687 | Glyma14g38580.1 | AT2G30490.1 | Encodes a cinnamate-4-hydroxylase. Mutations in this gene impact phenylpropanoid metabolism, growth and development. |
| GmaAffx.64204.1.S1_at | 0.04484 | -2.6155 | Glyma18g02450.2 | AT4G39640.2 | The gene encodes a gamma-glutamyltransferase (AKA gamma-glutamyl transpeptidase, EC 2.3.2.2) that is located in vascular tissues (predominantly phloem) of leaves and is involved in the degradation of glutathione. The encoded enzyme also mitigates oxidative stress by metabolizing GSSG (oxidized form of GSH - glutathione) in the apoplast. |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.92030.1.S1_ | 0.09531 | -2.5945 | Glyma19g44310.1 | AT2G46370.2 | | Encodes a jasmonate-amido synthetase that is a member of the GH3 family of proteins. JAR1 catalyzes the formation of a biologically active jasmonyl-isoleucine (JA-Ile) conjugate. JA-Ile promotes the interaction between JAZ1 and COI1 in the jasmonate signaling pathway. JAR1 localizes to the cytoplasm and is also a phytochrome A signaling component. JAR1 is an auxin-induced gene. Loss of function mutants are defective in a variety of responses to jasmonic acid. JAR1 has additional enzymatic activities in vitro, (e.g. the ability to synthesize adenosine 5'-tetraphosphate and other JA conjugates), but there are no data to show whether JAR1 catalyzes many of these reactions in vivo. JAR1 is involved in pathogen defense, sensitivity to ozone, and wound responses. |
| Gma.3504.2.S1_at | 0.04961 | -2.591 | Glyma13g00380.1 | AT4G31550.1 | | member of WRKY Transcription Factor; Group II-d; negative regulator of basal resistance to Pseudomonas syringae. |
| GmaAffx.65280.1.A1_ | 0.07824 | -2.5742 | Glyma13g26640.2 | AT3G27890.1 | | Encodes NAD(P)H:quinone reductase which is an FMN binding protein that catalyzes the reduction of quinone substrates to hydroquinones.The enzyme activity was confirmed by in vitro assay. |
| Gma.8020.3.S1_at | 0.07443 | -2.5206 | Glyma13g07490.1 | AT3G25780.1 | | Encodes allene oxide cyclase, one of the enzymes involved in jasmonic acid biosynthesis. One of four genes in Arabidopsis that encode this enzyme. mRNA expression is upregulated in senescing leaves. Note: Nomenclature for Arabidopsis allene oxide cyclase 3 (AOC3, AT3G25780) gene is based on Stenzel et al. 2003 Plant Molecular Biology 51:895-911. AOC3 (AT3G25780) is also referred to as AOC2 in He et al. 2002 Plant Physiology, 128:876-884. |
| Gma.338.1.S1_at | 0.02743 | -2.4379 | Glyma07g39020.1 | AT4G21960.1 | | Encodes AT4g21960 (AT4g21960/T8O5_170). |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.1584.1.S1_at | 0.08717 | -2.4168 | Glyma02g09550.2 | AT2G01830.1 | | Histidine kinase: cytokinin-binding receptor that transduces cytokinin signals across the plasma membrane |
| Gma.6889.1.S1_at | 0.04499 | -2.3871 | Glyma10g28900.1 | AT3G62550.1 | | universal stress protein (USP) family protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: response to stress; LOCATED IN: vacuole; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: UspA (InterPro:IPR006016), Rossmann-like alpha/beta/alpha sandwich fold (InterPro:IPR014729), Universal stress protein A (InterPro:IPR006015); BEST Arabidopsis thaliana protein match is: ethylene-responsive protein, putative (TAIR:AT1G09740.1); Has 1616 Blast hits to 1599 proteins in 385 species: Archae - 141; Bacteria - 962; Metazoa - 41; Fungi - 11; Plants - 403; Viruses - 0; Other Eukaryotes - 58 (source: NCBI BLink). |
| Gma.6889.1.S1_x_at | 0.05004 | -2.3802 | Glyma10g28900.1 | AT3G62550.1 | | universal stress protein (USP) family protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: response to stress; LOCATED IN: vacuole; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: UspA (InterPro:IPR006016), Rossmann-like alpha/beta/alpha sandwich fold (InterPro:IPR014729), Universal stress protein A (InterPro:IPR006015); BEST Arabidopsis thaliana protein match is: ethylene-responsive protein, putative (TAIR:AT1G09740.1); Has 1616 Blast hits to 1599 proteins in 385 species: Archae - 141; Bacteria - 962; Metazoa - 41; Fungi - 11; Plants - 403; Viruses - 0; Other Eukaryotes - 58 (source: NCBI BLink). |
| Gma.1502.1.S1_at | 0.08729 | -2.2853 | Glyma05g29400.1 | AT2G29420.1 | | Encodes glutathione transferase belonging to the tau class of GSTs. Naming convention according to Wagner et al. (2002). Induced by Salicylic acid. Independent of NPR1 for their induction by salicylic acid. |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.1510.2.A1_ | 0.08951 | -2.2765 | Glyma07g09090.1 | AT3G18030.1 | flavin mononucleotide flavoprotein involved in salt and osmotic tolerance HAL3A encodes for phosphopantothenoylcysteine decarboxylase being involved in Coenzyme A biosynthesis. HAL3A is predominant over another gene with the presumably same function (HAL3B). |
| Gma.10335.1.S1_at | 0.09725 | -2.2558 | Glyma13g27590.1 | AT4G21870.1 | 26.5 kDa class P-related heat shock protein (HSP26.5-P); INVOLVED IN: response to heat; LOCATED IN: cellular_component unknown; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Heat shock protein Hsp20 (InterPro:IPR002068), HSP20-like chaperone (InterPro:IPR008978); BEST Arabidopsis thaliana protein match is: HSP18.2 (heat shock protein 18.2) (TAIR:AT5G59720.1); Has 1539 Blast hits to 1539 proteins in 339 species: Archae - 25; Bacteria - 501; Metazoa - 0; Fungi - 42; Plants - 789; Viruses - 0; Other Eukaryotes - 182 (source: NCBI BLink). |
| Gma.2350.1.S1_at | 0.05394 | -2.2275 | No Soybean Match Id | AT5G60640.1 | Encodes a protein disulfide isomerase-like (PDIL) protein, a member of a multigene family within the thioredoxin (TRX) superfamily. Unlike several other PDI family members, transcript levels for this gene are not up-regulated in response to three different chemical inducers of ER stress (dithiothreitol, beta-mercaptoethanol, and tunicamycin). However, the level of transcripts for this gene is slightly elevated in atbzip60 mutants. |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.2009.1.S1_at | 0.03172 | -2.2145 | Glyma17g01550.3 | | AT1G11860.2 | aminomethyltransferase, putative; FUNCTIONS IN: aminomethyltransferase activity; INVOLVED IN: response to cadmium ion; LOCATED IN: in 6 components; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Glycine cleavage system T protein (InterPro:IPR006223), Glycine cleavage T-protein, N-terminal (InterPro:IPR006222), Glycine cleavage T-protein, C-terminal barrel (InterPro:IPR013977); BEST Arabidopsis thaliana protein match is: aminomethyltransferase (TAIR:AT1G60990.3); Has 16017 Blast hits to 16014 proteins in 1169 species: Archae - 69; Bacteria - 3430; Metazoa - 374; Fungi - 107; Plants - 67; Viruses - 0; Other Eukaryotes - 11970 (source: NCBI BLink). |
| GmaAffx.56846.1.S1_ | 0.07726 | -2.1753 | Glyma11g05160.1 | | AT3G49910.1 | 60S ribosomal protein L26 (RPL26A); FUNCTIONS IN: structural constituent of ribosome; INVOLVED IN: response to cold, translation; LOCATED IN: in 8 components; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Translation protein SH3-like (InterPro:IPR008991), Ribosomal protein L26, eukaryotic/archaeal (InterPro:IPR005756), Ribosomal protein L24, SH3-like (InterPro:IPR014723), KOW (InterPro:IPR005824), Ribosomal protein L24/L26, conserved site (InterPro:IPR005825); BEST Arabidopsis thaliana protein match is: 60S ribosomal protein L26 (RPL26B) (TAIR:AT5G67510.1); Has 870 Blast hits to 870 proteins in 315 species: Archae - 232; Bacteria - 6; Metazoa - 311; Fungi - 91; Plants - 68; Viruses - 0; Other Eukaryotes - 162 (source: NCBI BLink). |
| Gma.4207.1.S1_at | 0.02697 | -2.1469 | Glyma06g00630.1 | | AT4G34990.1 | Member of the R2R3 factor gene family. |

| | | | | | |
|---|---|---|---|---|---|
| Gma.8020.2.S1_a_at | 0.0845 | -2.1116 | No Soybean Match Id | AT1G13280.1 | Encodes allene oxide cyclase. One of four genes in Arabidopsis that encode this enzyme, which catalyzes an essential step in jasmonic acid biosynthesis. Gene expression is reduced during senescence, a process that involves jasmonic acid signalling pathway. |
| Gma.1445.2.S1_at | 0.05962 | -2.0966 | Glyma15g00600.2 | AT1G15950.1 | Encodes a cinnamoyl CoA reductase. Involved in lignin biosynthesis. |
| Gma.7487.1.S1_at | 0.03315 | -2.0808 | Glyma04g09310.1 | AT5G49630.1 | Is a high affinity amino acid transporter capable of transporting aspartate and tryptophan. May be involved in the amino acid uptake from xylem. |
| GmaAffx.6478.1.S1_s | 0.06847 | -2.0507 | Glyma13g00380.1 | AT4G31550.1 | member of WRKY Transcription Factor; Group II-d; negative regulator of basal resistance to Pseudomonas syringae. |
| GmaAffx.68100.1.S1_ | 0.03377 | -2.0354 | Glyma04g30350.1 | AT5G17920.2 | Encodes a cytosolic cobalamin-independent methionine synthase, involved in methionine regeneration via the activated methyl cycle (SAM cycle). The protein undergoes thiolation following treatment with the oxidant tert-butylhydroperoxide. |
| Gma.5854.1.S1_at | 0.09622 | -2.0027 | Glyma08g05640.1 | AT3G18060.1 | transducin family protein / WD-40 repeat family protein; FUNCTIONS IN: nucleotide binding; INVOLVED IN: biological_process unknown; LOCATED IN: CUL4 RING ubiquitin ligase complex; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: WD40 repeat-like (InterPro:IPR011046), WD40 repeat, region (InterPro:IPR017986), WD40/YVTN repeat-like (InterPro:IPR015943), WD40 repeat (InterPro:IPR001680), Nitrous oxide reductase, N-terminal (InterPro:IPR011045); BEST Arabidopsis thaliana protein match is: transducin family protein / WD-40 repeat family protein (TAIR:AT2G01330.1); Has 34874 Blast hits to 15896 proteins in 540 species: Archae - 42; Bacteria - 5137; Metazoa - 15525; Fungi - 6232; Plants - 3080; Viruses - 0; Other Eukaryotes - 4858 (source: NCBI BLink). |

| GmaAffx.15471.1.S1_ | 0.07811 | -2.0004 | Glyma12g07110.2 | AT5G59780.2 | Encodes a putative transcription factor (MYB59). |
|---|---|---|---|---|---|
| Gma.6384.1.S1_at | 0.05171 | -1.9952 | Glyma12g08650.1 | AT2G28190.1 | Encodes a chloroplastic copper/zinc superoxide dismutase CSD2 that can detoxify superoxide radicals. Its expression is affected by miR398-directed mRNA cleavage. |
| Gma.7270.1.S1_a_at | 0.05952 | -1.9779 | No Soybean Match Id | AT3G29320.1 | Encodes a plastidic alpha-glucan phosphorylase. In vitro, the enzyme has a preference for maltooligosaccharides, such as maltoheptaose. |
| Gma.3504.1.S1_at | 0.09471 | -1.965 | Glyma17g06450.1 | AT4G31550.1 | member of WRKY Transcription Factor; Group II-d; negative regulator of basal resistance to Pseudomonas syringae. |
| Gma.4189.1.S1_at | 0.05361 | -1.9415 | Glyma17g01720.1 | AT4G21960.1 | Encodes AT4g21960 (AT4g21960/T8O5_170). |
| Gma.17711.1.S1_at | 0.03392 | -1.9373 | Glyma03g40640.1 | AT3G04790.1 | ribose 5-phosphate isomerase-related; FUNCTIONS IN: ribose-5-phosphate isomerase activity; INVOLVED IN: defense response to bacterium, reductive pentose-phosphate cycle; LOCATED IN: thylakoid, chloroplast thylakoid membrane, chloroplast stroma, chloroplast, chloroplast envelope; EXPRESSED IN: 26 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Ribose 5-phosphate isomerase (InterPro:IPR004788); BEST Arabidopsis thaliana protein match is: ribose-5-phosphate isomerase (TAIR:AT2G01290.1); Has 3164 Blast hits to 3164 proteins in 1122 species: Archae - 153; Bacteria - 1911; Metazoa - 90; Fungi - 101; Plants - 84; Viruses - 0; Other Eukaryotes - 825 (source: NCBI BLink). |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.91273.1.S1_ | 0.0958 | -1.9322 | Glyma06g45370.1 | AT4G10270.1 | | wound-responsive family protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: response to wounding; EXPRESSED IN: 20 plant structures; EXPRESSED DURING: 11 growth stages; BEST Arabidopsis thaliana protein match is: wound-responsive protein, putative (TAIR:AT4G10265.1); Has 110 Blast hits to 109 proteins in 12 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 110; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| Gma.3504.2.S1_a_at | 0.07886 | -1.9234 | Glyma13g00380.1 | AT4G31550.1 | | member of WRKY Transcription Factor; Group II-d; negative regulator of basal resistance to Pseudomonas syringae. |
| Gma.3145.1.S1_at | 0.07243 | -1.9011 | Glyma19g05120.2 | AT3G02360.1 | | 6-phosphogluconate dehydrogenase family protein; FUNCTIONS IN: in 6 functions; INVOLVED IN: response to salt stress; LOCATED IN: peroxisome; EXPRESSED IN: 27 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: 6-phosphogluconate dehydrogenase, NAD-binding (InterPro:IPR006115), 6-phosphogluconate dehydrogenase, C-terminal-like (InterPro:IPR008927), Dehydrogenase, multihelical (InterPro:IPR013328), 6-phosphogluconate dehydrogenase, decarboxylating (InterPro:IPR006113), 6-phosphogluconate dehydrogenase, C-terminal (InterPro:IPR006114), 6-phosphogluconate dehydrogenase (InterPro:IPR006183), NAD(P)-binding (InterPro:IPR016040), Fibritin/6-phosphogluconate dehydrogenase, C-terminal extension (InterPro:IPR012284); BEST Arabidopsis thaliana protein match is: 6-phosphogluconate dehydrogenase family protein (TAIR:AT5G41670.2); Has 8125 Blast hits to 8086 proteins in 1474 species: Archae - 54; Bacteria - 4368; Metazoa - 506; Fungi - 152; Plants - 199; Viruses - 2; Other Eukaryotes - 2844 (source: NCBI BLink). |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.17882.1.S1_at | 0.0774 | -1.8962 | Glyma15g01990.1 | AT1G52200.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: response to oxidative stress; LOCATED IN: plasma membrane; EXPRESSED IN: 10 plant structures; EXPRESSED DURING: 6 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function Cys-rich (InterPro:IPR006461); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT3G18470.1); Has 371 Blast hits to 370 proteins in 48 species: Archae - 0; Bacteria - 0; Metazoa - 28; Fungi - 47; Plants - 295; Viruses - 0; Other Eukaryotes - 1 (source: NCBI BLink). |
| Gma.6515.1.S1_s_at | 0.05952 | -1.8641 | Glyma02g46050.2 | AT3G20390.1 | | endoribonuclease L-PSP family protein; FUNCTIONS IN: endoribonuclease activity; INVOLVED IN: response to cadmium ion; LOCATED IN: thylakoid, mitochondrion, chloroplast, plastid, vacuole; EXPRESSED IN: 27 plant structures; EXPRESSED DURING: 16 growth stages; CONTAINS InterPro DOMAIN/s: Endoribonuclease L-PSP (InterPro:IPR006175), Endoribonuclease L-PSP/chorismate mutase-like (InterPro:IPR013813), YjgF-like protein (InterPro:IPR006056); Has 6124 Blast hits to 6041 proteins in 1217 species: Archae - 83; Bacteria - 4022; Metazoa - 194; Fungi - 247; Plants - 42; Viruses - 0; Other Eukaryotes - 1536 (source: NCBI BLink). |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.48466.1.S1_ | 0.09403 | -1.8049 | Glyma10g41450.2 | AT2G01140.1 | | fructose-bisphosphate aldolase, putative; FUNCTIONS IN: fructose-bisphosphate aldolase activity; INVOLVED IN: response to oxidative stress, response to cadmium ion, pentose-phosphate shunt; LOCATED IN: mitochondrion, chloroplast, plastoglobule; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Aldolase-type TIM barrel (InterPro:IPR013785), Fructose-bisphosphate aldolase, class-I (InterPro:IPR000741); BEST Arabidopsis thaliana protein match is: fructose-bisphosphate aldolase, putative (TAIR:AT4G38970.1); Has 4200 Blast hits to 4196 proteins in 676 species: Archae - 0; Bacteria - 431; Metazoa - 1018; Fungi - 2; Plants - 339; Viruses - 0; Other Eukaryotes - 2410 (source: NCBI BLink). |
| Gma.8489.1.S1_at | 0.09529 | -1.7933 | Glyma12g16380.2 | AT5G54770.1 | | Encodes a thiamine biosynthetic gene that has a dual function in thiamine biosynthesis and mitochondrial DNA damage tolerance. It appears to be involved in producing the thiazole portion of thiamine (vitamin B1). A crystal structure of the protein reveals that it forms a 2-ring homo-octamer. |
| Gma.3841.1.S1_at | 0.06412 | -1.7925 | Glyma08g42820.2 | AT3G20390.1 | | endoribonuclease L-PSP family protein; FUNCTIONS IN: endoribonuclease activity; INVOLVED IN: response to cadmium ion; LOCATED IN: thylakoid, mitochondrion, chloroplast, plastid, vacuole; EXPRESSED IN: 27 plant structures; EXPRESSED DURING: 16 growth stages; CONTAINS InterPro DOMAIN/s: Endoribonuclease L-PSP (InterPro:IPR006175), Endoribonuclease L-PSP/chorismate mutase-like (InterPro:IPR013813), YjgF-like protein (InterPro:IPR006056); Has 6124 Blast hits to 6041 proteins in 1217 species: Archae - 83; Bacteria - 4022; Metazoa - 194; Fungi - 247; Plants - 42; Viruses - 0; Other Eukaryotes - 1536 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.54278.1.S1_ | 0.07771 | -1.7824 | Glyma13g03600.1 | AT1G21750.1 | Encodes a protein disulfide isomerase-like (PDIL) protein, a member of a multigene family within the thioredoxin (TRX) superfamily; isoform contains non-consensus GA donor splice site at intron 9. Transcript levels for this gene are up-regulated in response to three different chemical inducers of ER stress (dithiothreitol, beta-mercaptoethanol, and tunicamycin). Neither AtIRE1-2 nor AtbZIP60 appear to be required for this response. |
| Gma.15478.1.S1_at | 0.09671 | -1.7745 | No Soybean Match Id | AT1G02500.2 | encodes a S-adenosylmethionine synthetase. SAM1 is regulated by protein S-nitrosylation. The covalent binding of nitric oxide (NO) to the Cys114 residue inhibits the enzyme activity. |
| Gma.11321.1.S1_s_a | 0.05603 | -1.7699 | Glyma03g25200.1 | AT4G17670.1 | senescence-associated protein-related; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF581 (InterPro:IPR007650); BEST Arabidopsis thaliana protein match is: senescence-associated protein-related (TAIR:AT5G47060.1); Has 285 Blast hits to 285 proteins in 14 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 285; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| Gma.3920.1.S1_at | 0.06865 | -1.7688 | No Soybean Match Id | AT5G17920.2 | Encodes a cytosolic cobalamin-independent methionine synthase, involved in methionine regeneration via the activated methyl cycle (SAM cycle). The protein undergoes thiolation following treatment with the oxidant tert-butylhydroperoxide. |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.7492.1.S1_at | 0.07755 | -1.7658 | Glyma12g02510.2 | | AT3G22200.1 | Genetically redundant with POP3;mediates pollen tube guidance. Double mutants are self sterile; gamma-aminobutyrate transaminase subunit precursor; nuclear gene for mitochondrial product. Encodes gamma-aminobutyrate transaminase that uses pyruvate instead of alpha-ketoglutarate as cosubstrate. Mutations in POP2/HER1 render roots resistant to the inhibitory growth effects of the volatile organic compound E-2-hexenal implicated in plant defense. |
| Gma.5533.1.S1_at | 0.03473 | -1.7648 | Glyma20g33060.1 | | AT3G52990.1 | pyruvate kinase, putative; FUNCTIONS IN: pyruvate kinase activity, potassium ion binding, magnesium ion binding, catalytic activity; INVOLVED IN: response to cadmium ion, glycolysis; LOCATED IN: membrane; EXPRESSED IN: guard cell, cultured cell, leaf; EXPRESSED DURING: seedling growth; CONTAINS InterPro DOMAIN/s: Pyruvate/Phosphoenolpyruvate kinase, catalytic core (InterPro:IPR015813), Pyruvate kinase, alpha/beta (InterPro:IPR015794), Pyruvate kinase, beta-barrel-like (InterPro:IPR011037), Pyruvate kinase (InterPro:IPR001697), Pyruvate kinase, C-terminal-like (InterPro:IPR015795), Pyruvate kinase, barrel (InterPro:IPR015793); BEST Arabidopsis thaliana protein match is: pyruvate kinase, putative (TAIR:AT2G36580.1); Has 6549 Blast hits to 6522 proteins in 1516 species: Archae - 99; Bacteria - 3191; Metazoa - 483; Fungi - 169; Plants - 286; Viruses - 0; Other Eukaryotes - 2321 (source: NCBI BLink). |
| Gma.10631.1.S1_at | 0.09193 | -1.7638 | Glyma06g18470.1 | | AT4G24770.1 | Encodes a chloroplast RNA-binding protein. A substrate of the type III effector HopU1 (mono-ADP-ribosyltransferase). Required for editing and stability of specific chloroplast mRNAs. |

| | | | | | |
|---|---|---|---|---|---|
| Gma.2677.1.S1_s_at | 0.06412 | -1.7308 | Glyma19g44310.1 | AT2G46370.2 | Encodes a jasmonate-amido synthetase that is a member of the GH3 family of proteins. JAR1 catalyzes the formation of a biologically active jasmonyl-isoleucine (JA-Ile) conjugate. JA-Ile promotes the interaction between JAZ1 and COI1 in the jasmonate signaling pathway. JAR1 localizes to the cytoplasm and is also a phytochrome A signaling component. JAR1 is an auxin-induced gene. Loss of function mutants are defective in a variety of responses to jasmonic acid. JAR1 has additional enzymatic activities in vitro, (e.g. the ability to synthesize adenosine 5'-tetraphosphate and other JA conjugates), but there are no data to show whether JAR1 catalyzes many of these reactions in vivo. JAR1 is involved in pathogen defense, sensitivity to ozone, and wound responses. |
| Gma.2773.2.S1_at | 0.06716 | -1.7209 | Glyma19g33330.1 | AT1G65870.1 | disease resistance-responsive family protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: lignan biosynthetic process, defense response; LOCATED IN: cell wall; EXPRESSED IN: 13 plant structures; EXPRESSED DURING: 9 growth stages; CONTAINS InterPro DOMAIN/s: Plant disease resistance response protein (InterPro:IPR004265); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT1G22900.1); Has 517 Blast hits to 516 proteins in 34 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 517; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | | | | dehydration-responsive family protein; INVOLVED IN: biological_process unknown; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF248, methyltransferase putative (InterPro:IPR004159); BEST Arabidopsis thaliana protein match is: dehydration-responsive protein-related (TAIR:AT5G06050.1); Has 1179 Blast hits to 1086 proteins in 154 species: Archae - 2; Bacteria - 119; Metazoa - 231; Fungi - 82; Plants - 566; Viruses - 29; Other Eukaryotes - 150 (source: NCBI BLink). |
| GmaAffx.85050.1.S1_ | 0.04364 | -1.7169 | Glyma02g41770.1 | AT2G39750.1 | | |
| Gma.3253.1.S1_at | 0.05811 | -1.7056 | Glyma05g04520.1 | AT4G17260.1 | L-lactate dehydrogenase, putative; FUNCTIONS IN: in 6 functions; INVOLVED IN: response to salt stress, response to abscisic acid stimulus; LOCATED IN: plasma membrane; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: L-lactate dehydrogenase (InterPro:IPR011304), NAD(P)-binding (InterPro:IPR016040), L-lactate/malate dehydrogenase (InterPro:IPR001557), Lactate/malate dehydrogenase (InterPro:IPR001236), L-lactate dehydrogenase, active site (InterPro:IPR018177), Lactate dehydrogenase/glycoside hydrolase, family 4, C-terminal (InterPro:IPR015955); BEST Arabidopsis thaliana protein match is: PMDH2 (peroxisomal NAD-malate dehydrogenase 2); malate dehydrogenase (TAIR:AT5G09660.2); Has 9812 Blast hits to 9805 proteins in 2184 species: Archae - 185; Bacteria - 5417; Metazoa - 1010; Fungi - 354; Plants - 275; Viruses - 0; Other Eukaryotes - 2571 (source: NCBI BLink). |
| Gma.350.1.S1_at | 0.08135 | -1.6997 | Glyma20g09810.1 | AT3G14940.1 | One of four genes encoding phosphoenolpyruvate carboxylase, its mRNA is most abundantly expressed in roots and siliques. | |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.2749.1.S1_at | 0.04759 | -1.6985 | Glyma10g40140.1 | AT1G80600.1 | | Encodes HopW1-1-Interacting protein 1 (WIN1). Interacts with the P. syringae effector HopW1-1. WIN1 is a putative acetylornithine transaminase. Modulates plant defenses against bacteria. Three WIN proteins are identified so far (WIN1: AT1G80600; WIN2: AT4G31750; WIN3: AT5G13320). |
| GmaAffx.6552.1.A1_ | 0.08318 | -1.6937 | Glyma02g35780.1 | AT3G11630.1 | | Encodes a 2-Cys peroxiredoxin (2-Cys PrxA) that contains two catalytic Cys residues. |
| GmaAffx.43676.1.S1_ | 0.09985 | -1.6868 | Glyma10g02040.1 | AT2G47510.2 | | fumarase (FUM1) |
| GmaAffx.93074.1.S1_ | 0.07978 | -1.6611 | No Soybean Match Id | AT3G02360.1 | | 6-phosphogluconate dehydrogenase family protein; FUNCTIONS IN: in 6 functions; INVOLVED IN: response to salt stress; LOCATED IN: peroxisome; EXPRESSED IN: 27 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: 6-phosphogluconate dehydrogenase, NAD-binding (InterPro:IPR006115), 6-phosphogluconate dehydrogenase, C-terminal-like (InterPro:IPR008927), Dehydrogenase, multihelical (InterPro:IPR013328), 6-phosphogluconate dehydrogenase, decarboxylating (InterPro:IPR006113), 6-phosphogluconate dehydrogenase, C-terminal (InterPro:IPR006114), 6-phosphogluconate dehydrogenase (InterPro:IPR006183), NAD(P)-binding (InterPro:IPR016040), Fibritin/6-phosphogluconate dehydrogenase, C-terminal extension (InterPro:IPR012284); BEST Arabidopsis thaliana protein match is: 6-phosphogluconate dehydrogenase family protein (TAIR:AT5G41670.2); Has 8125 Blast hits to 8086 proteins in 1474 species: Archae - 54; Bacteria - 4368; Metazoa - 506; Fungi - 152; Plants - 199; Viruses - 2; Other Eukaryotes - 2844 (source: NCBI BLink). |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.2350.1.S1_s_at | 0.09671 | -1.6453 | Glyma13g40130.1 | AT5G60640.1 | | Encodes a protein disulfide isomerase-like (PDIL) protein, a member of a multigene family within the thioredoxin (TRX) superfamily. Unlike several other PDI family members, transcript levels for this gene are not up-regulated in response to three different chemical inducers of ER stress (dithiothreitol, beta-mercaptoethanol, and tunicamycin). However, the level of transcripts for this gene is slightly elevated in atbzip60 mutants. |
| Gma.4312.1.S1_at | 0.0982 | -1.633 | Glyma08g05200.1 | AT2G31570.1 | | glutathione peroxidase GPx |
| GmaAffx.89649.1.S1_ | 0.05226 | -1.6312 | Glyma05g34490.4 | AT2G43350.1 | | Glutathione peroxidase. Functions as both a redox transducer and a scavenger in abscisic acid and drought stress responses. Interacts with ABI2 and ABI1. |
| Gma.2139.2.S1_s_at | 0.04164 | -1.6173 | Glyma13g41960.1 | AT3G59480.1 | | pfkB-type carbohydrate kinase family protein; FUNCTIONS IN: kinase activity, ribokinase activity; INVOLVED IN: D-ribose metabolic process, acetate fermentation, sucrose biosynthetic process, sucrose catabolic process, using beta-fructofuranosidase; LOCATED IN: cellular_component unknown; EXPRESSED IN: 9 plant structures; EXPRESSED DURING: 4 anthesis, C globular stage, petal differentiation and expansion stage; CONTAINS InterPro DOMAIN/s: Carbohydrate/purine kinase (InterPro:IPR011611), Ribokinase (InterPro:IPR002139), Carbohydrate/puine kinase, PfkB, conserved site (InterPro:IPR002173); BEST Arabidopsis thaliana protein match is: pfkB-type carbohydrate kinase family protein (TAIR:AT2G31390.1); Has 12101 Blast hits to 12099 proteins in 1317 species: Archae - 211; Bacteria - 7841; Metazoa - 113; Fungi - 101; Plants - 260; Viruses - 0; Other Eukaryotes - 3575 (source: NCBI BLink). |
| Gma.16202.1.S1_at | 0.05952 | -1.6167 | Glyma15g04670.3 | AT2G39460.1 | | Encodes a 60S ribosomal protein L23aA (AtrpL23aA). Paralog of RLPL23aB. |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.8712.1.S1_s | 0.05811 | -1.6115 | No Soybean Match Id | AT5G05340.1 | | peroxidase, putative; FUNCTIONS IN: electron carrier activity, peroxidase activity, heme binding; INVOLVED IN: response to oxidative stress; LOCATED IN: apoplast, cell wall; EXPRESSED IN: 9 plant structures; EXPRESSED DURING: 4 anthesis, LP.04 four leaves visible, C globular stage, petal differentiation and expansion stage; CONTAINS InterPro DOMAIN/s: Haem peroxidase (InterPro:IPR010255), Plant peroxidase (InterPro:IPR000823), Haem peroxidase, plant/fungal/bacterial (InterPro:IPR002016); BEST Arabidopsis thaliana protein match is: peroxidase, putative (TAIR:AT5G58400.1); Has 2945 Blast hits to 2930 proteins in 223 species: Archae - 0; Bacteria - 4; Metazoa - 2; Fungi - 110; Plants - 2778; Viruses - 0; Other Eukaryotes - 51 (source: NCBI BLink). |
| GmaAffx.84808.1.S1_ | 0.03908 | -1.5903 | Glyma13g20810.2 | AT5G03280.1 | | Involved in ethylene signal transduction. Acts downstream of CTR1. Positively regulates ORE1 and negatively regulates mir164A,B,C to regulate leaf senescence. |
| Gma.5111.2.S1_at | 0.09403 | -1.5836 | Glyma05g03120.1 | AT5G63400.1 | | encodes a protein similar to adenylate kinase. |
| GmaAffx.80986.1.S1_ | 0.07755 | -1.5822 | Glyma17g10990.2 | AT5G61410.2 | | Arabidopsis thaliana ribulose-5-phosphate-3-epimerase mRNA |
| Gma.5111.1.S1_at | 0.07766 | -1.5577 | No Soybean Match Id | AT5G63400.1 | | encodes a protein similar to adenylate kinase. |
| GmaAffx.85352.1.S1_ | 0.09148 | -1.5424 | Glyma03g33850.1 | AT5G03280.1 | | Involved in ethylene signal transduction. Acts downstream of CTR1. Positively regulates ORE1 and negatively regulates mir164A,B,C to regulate leaf senescence. |
| Gma.3622.1.S1_at | 0.07771 | -1.5416 | Ambiguous Hit | AT5G20720.2 | | Encodes a chloroplast co-chaperonin with similarity to CPN21 from spinach, E.coli GroES. |
| GmaAffx.5378.1.S1_a | 0.06392 | -1.5302 | Glyma11g36580.1 | AT4G13850.2 | | Encodes a glycine-rich RNA-binding protein. Gene expression is induced by cold. |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.459.1.S1_s_ | 0.08446 | -1.5118 | Glyma06g00400.1 | AT5G17690.1 | | Regulates the meristem response to light signals and the maintenance of inflorescence meristem identity. Influences developmental processes controlled by APETALA1. TFL2 silences specific genes within euchromatin but not genes positioned in heterochromatin. TFL2 protein localized preferentially to euchromatic regions and not to heterochromatic chromocenters. Involved in euchromatin organization. Required for epigenetic maintenance of the vernalized state. |
| Gma.2162.1.S1_at | 0.05952 | -1.4901 | Glyma20g25790.1 | AT2G01140.1 | | fructose-bisphosphate aldolase, putative; FUNCTIONS IN: fructose-bisphosphate aldolase activity; INVOLVED IN: response to oxidative stress, response to cadmium ion, pentose-phosphate shunt; LOCATED IN: mitochondrion, chloroplast, plastoglobule; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Aldolase-type TIM barrel (InterPro:IPR013785), Fructose-bisphosphate aldolase, class-I (InterPro:IPR000741); BEST Arabidopsis thaliana protein match is: fructose-bisphosphate aldolase, putative (TAIR:AT4G38970.1); Has 4200 Blast hits to 4196 proteins in 676 species: Archae - 0; Bacteria - 431; Metazoa - 1018; Fungi - 2; Plants - 339; Viruses - 0; Other Eukaryotes - 2410 (source: NCBI BLink). |
| GmaAffx.37206.1.S1_ | 0.0958 | -1.4782 | Glyma07g30210.1 | AT2G14170.1 | | Arabidopsis thaliana methylmalonate-semialdehyde dehydrogenase |
| Gma.6719.1.S1_at | 0.07219 | -1.4719 | Glyma16g10730.1 | AT3G25230.1 | | Encodes a a high molecular weight member of the FK506 binding protein (FKBP) family. It has three FKBP12-like domains, tetratricopeptide repeats, and a putative calmodulin binding domain. Modulates thermotolerance by interacting with HSP90.1 and affecting the accumulation of HsfA2-regulated sHSPs. |
| Gma.9110.1.S1_at | 0.06264 | -1.4645 | Glyma13g32660.1 | AT5G09650.1 | | Encodes a protein with inorganic pyrophosphatase activity. |

| | | | | | |
|---|---|---|---|---|---|
| Gma.4312.1.S1_x_at | 0.09726 | -1.4613 | Glyma08g05200.1 | AT2G31570.1 | glutathione peroxidase GPx |
| Gma.15462.3.S1_at | 0.05226 | -1.459 | Glyma10g06620.1 | AT2G37220.1 | Encodes a chloroplast RNA binding protein. A substrate of the type III effector HopU1 (mono-ADP-ribosyltransferase). Protein is tyrosine-phosphorylated and its phosphorylation state is modulated in response to ABA in Arabidopsis thaliana seeds. |
| Gma.3301.1.S1_at | 0.09251 | -1.4262 | Glyma06g00440.1 | AT4G02600.2 | A member of a large family of seven-transmembrane domain proteins specific to plants, homologs of the barley mildew resistance locus o (MLO) protein. The Arabidopsis genome contains 15 genes encoding MLO proteins, with localization in plasma membrane. Phylogenetic analysis revealed four clades of closely-related AtMLO genes. ATMLO1 belongs to the clade II, with ATMLO13 and ATMLO15. The gene is expressed during early seedling growth, in root and cotyledon vascular system, in pollen and in papillae, as shown by GUS activity patterns. The expression of several phylogenetically closely-related AtMLO genes showed similar or overlapping tissue specificity and analogous responsiveness to external stimuli, suggesting functional redundancy, co-function, or antagonistic function(s). |
| GmaAffx.92104.1.S1 | 0.06238 | -1.4217 | No Soybean Match Id | AT3G13920.3 | eukaryotic translation initiation factor 4A-1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.20940.1.S1_ | 0.03369 | -1.4162 | Glyma17g05840.1 | AT2G31670.1 | | FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: peroxisome, chloroplast stroma, chloroplast; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 16 growth stages; CONTAINS InterPro DOMAIN/s: Stress responsive alpha-beta barrel (InterPro:IPR013097), Dimeric alpha-beta barrel (InterPro:IPR011008); BEST Arabidopsis thaliana protein match is: DABB1 (DIMERIC A/B BARREL DOMAINS-PROTEIN 1) (TAIR:AT1G51360.1); Has 138 Blast hits to 135 proteins in 38 species: Archae - 0; Bacteria - 51; Metazoa - 0; Fungi - 4; Plants - 74; Viruses - 0; Other Eukaryotes - 9 (source: NCBI BLink). |
| Gma.7270.1.S1_at | 0.09346 | -1.4093 | Glyma13g06040.1 | AT3G29320.1 | | Encodes a plastidic alpha-glucan phosphorylase. In vitro, the enzyme has a preference for maltooligosaccharides, such as maltoheptaose. |
| Gma.3582.1.S1_at | 0.07267 | -1.3697 | Glyma17g18500.1 | AT3G50210.3 | | 2-oxoacid-dependent oxidase, putative; FUNCTIONS IN: oxidoreductase activity, iron ion binding; INVOLVED IN: aging, cellular response to starvation; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Isopenicillin N synthase (InterPro:IPR002283), 2OG-Fe(II) oxygenase (InterPro:IPR005123); BEST Arabidopsis thaliana protein match is: DIN11 (DARK INDUCIBLE 11); iron ion binding / oxidoreductase (TAIR:AT3G49620.1); Has 6172 Blast hits to 6127 proteins in 685 species: Archae - 0; Bacteria - 736; Metazoa - 131; Fungi - 690; Plants - 2990; Viruses - 0; Other Eukaryotes - 1625 (source: NCBI BLink). |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.90703.1.S1_ | 0.07717 | -1.3677 | Glyma06g15030.1 | AT5G05340.1 | | peroxidase, putative; FUNCTIONS IN: electron carrier activity, peroxidase activity, heme binding; INVOLVED IN: response to oxidative stress; LOCATED IN: apoplast, cell wall; EXPRESSED IN: 9 plant structures; EXPRESSED DURING: 4 anthesis, LP.04 four leaves visible, C globular stage, petal differentiation and expansion stage; CONTAINS InterPro DOMAIN/s: Haem peroxidase (InterPro:IPR010255), Plant peroxidase (InterPro:IPR000823), Haem peroxidase, plant/fungal/bacterial (InterPro:IPR002016); BEST Arabidopsis thaliana protein match is: peroxidase, putative (TAIR:AT5G58400.1); Has 2945 Blast hits to 2930 proteins in 223 species: Archae - 0; Bacteria - 4; Metazoa - 2; Fungi - 110; Plants - 2778; Viruses - 0; Other Eukaryotes - 51 (source: NCBI BLink). |
| GmaAffx.90444.1.S1_ | 0.05811 | -1.3225 | Glyma11g21260.1 | AT3G27890.1 | | Encodes NAD(P)H:quinone reductase which is an FMN binding protein that catalyzes the reduction of quinone substrates to hydroquinones.The enzyme activity was confirmed by in vitro assay. |
| Gma.17715.1.S1_at | 0.06847 | -1.3087 | Glyma13g27220.4 | AT3G63520.1 | | Encodes a protein with 9-cis-epoxycarotenoid dioxygenase activity. The enzyme was shown to act on a variety of carotenoid including β-carotene, lutein, zeaxanthin, and all-trans-violaxanthin. When those compounds are used as substrates, the major reaction product detected is a C14 dialdehyde: 4,9-dimethyldodeca-2,4,6,8,10-pentaene-1,12-dial. The enzyme did not cleave as efficiently carotenoids containing 9-cis-double or allenic bonds. |
| GmaAffx.86228.1.S1_ | 0.09403 | -1.2629 | Glyma07g09240.1 | AT1G65980.1 | | thioredoxin-dependent peroxidase |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.49231.1.S1_ | 0.08446 | -1.231 | Glyma08g09730.1 | AT2G01980.1 | Encodes a plasma membrane-localized Na+/H+ antiporter SOS1. Functions in the extrusion of toxic Na+ from cells and is essential for plant salt tolerance. Has 12 predicted transmembrane domains in the N-terminal region and a long cytoplasmic tail of approx. 700 aa at the C-terminal side. SOS1 interacts through its predicted cytoplasmic tail with RCD1, a regulator of oxidative-stress responses, suggesting that SOS1 might function in oxidative-stress tolerance. |

Table 9 Candidate genes related to transcription regulation

| Affy ID | Q-value | Fold change | Soybean Gene | Arabidopsis homolog | Description |
|---|---|---|---|---|---|
| Gma.8113.1.A1_at | 0.026789 | 15.77957 | Glyma15g08480.2 | AT4G28500.1 | ARABIDOPSIS NAC DOMAIN CONTAINING PROTEIN 73 (ANAC073); FUNCTIONS IN: transcription factor activity, transcription activator activity; INVOLVED IN: multicellular organismal development, regulation of secondary cell wall thickening; LOCATED IN: nucleus; EXPRESSED IN: 9 plant structures; EXPRESSED DURING: 4 anthesis, petal differentiation and expansion stage; CONTAINS InterPro DOMAIN/s: No apical meristem (NAM) protein (InterPro:IPR003441); BEST Arabidopsis thaliana protein match is: ANAC010 (ARABIDOPSIS NAC DOMAIN CONTAINING PROTEIN 10); transcription activator/ transcription factor (TAIR:AT1G28470.1); Has 1219 Blast hits to 1217 proteins in 49 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 1218; Viruses - 0; Other Eukaryotes - 1 (source: NCBI BLink). |
| GmaAffx.50811.2.S1_ | 0.057908 | 11.47505 | Glyma15g08480.2 | AT4G28500.1 | ARABIDOPSIS NAC DOMAIN CONTAINING PROTEIN 73 (ANAC073); FUNCTIONS IN: transcription factor activity, transcription activator activity; INVOLVED IN: multicellular organismal development, regulation of secondary cell wall thickening; LOCATED IN: nucleus; EXPRESSED IN: 9 plant structures; EXPRESSED DURING: 4 anthesis, petal differentiation and expansion stage; CONTAINS InterPro DOMAIN/s: No apical meristem (NAM) protein (InterPro:IPR003441); BEST Arabidopsis thaliana protein match is: ANAC010 (ARABIDOPSIS NAC DOMAIN CONTAINING PROTEIN 10); transcription activator/ transcription factor (TAIR:AT1G28470.1); Has 1219 Blast hits to 1217 proteins in 49 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 1218; Viruses - 0; Other Eukaryotes - 1 (source: NCBI BLink). |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.986.1.S1_at | 0.033752 | 11.08897 | Glyma15g04570.1 | AT2G28710.1 | | zinc finger (C2H2 type) family protein; FUNCTIONS IN: transcription factor activity, zinc ion binding, nucleic acid binding; INVOLVED IN: regulation of transcription; LOCATED IN: intracellular; EXPRESSED IN: 8 plant structures; EXPRESSED DURING: 4 anthesis; CONTAINS InterPro DOMAIN/s: Zinc finger, C2H2-like (InterPro:IPR015880), Zinc finger, C2H2-type (InterPro:IPR007087); BEST Arabidopsis thaliana protein match is: RHL41 (RESPONSIVE TO HIGH LIGHT 41); nucleic acid binding / transcription factor/ zinc ion binding (TAIR:AT5G59820.1); Has 907 Blast hits to 871 proteins in 97 species: Archae - 0; Bacteria - 0; Metazoa - 429; Fungi - 0; Plants - 469; Viruses - 0; Other Eukaryotes - 9 (source: NCBI BLink). |
| GmaAffx.65341.1.A1_ | 0.079456 | 10.01211 | Glyma08g14630.1 | AT5G10120.1 | | ethylene insensitive 3 family protein; FUNCTIONS IN: transcription factor activity; INVOLVED IN: regulation of transcription; LOCATED IN: nucleus; EXPRESSED IN: stem, hypocotyl, flower, root, seed; EXPRESSED DURING: F mature embryo stage, petal differentiation and expansion stage; CONTAINS InterPro DOMAIN/s: Ethylene insensitive 3 (InterPro:IPR006957); BEST Arabidopsis thaliana protein match is: ethylene insensitive 3 family protein (TAIR:AT5G65100.1); Has 248 Blast hits to 233 proteins in 35 species: Archae - 0; Bacteria - 0; Metazoa - 3; Fungi - 0; Plants - 237; Viruses - 0; Other Eukaryotes - 8 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.17736.1.S1_at | 0.053776 | 7.53273 | Glyma03g33070.1 | AT2G37430.1 | zinc finger (C2H2 type) family protein (ZAT11); FUNCTIONS IN: transcription factor activity, zinc ion binding, nucleic acid binding; INVOLVED IN: response to chitin, regulation of transcription; LOCATED IN: intracellular; EXPRESSED IN: cotyledon, root, petiole; CONTAINS InterPro DOMAIN/s: Zinc finger, C2H2-like (InterPro:IPR015880), Zinc finger, C2H2-type (InterPro:IPR007087); BEST Arabidopsis thaliana protein match is: zinc finger (C2H2 type) family protein (TAIR:AT3G53600.1); Has 4550 Blast hits to 3560 proteins in 108 species: Archae - 0; Bacteria - 0; Metazoa - 3995; Fungi - 10; Plants - 464; Viruses - 0; Other Eukaryotes - 81 (source: NCBI BLink). |
| GmaAffx.41136.1.S1_ | 0.06315 | 5.734763 | No Soybean Matc | AT2G34710.1 | Dominant PHB mutations cause transformation of abaxial leaf fates into adaxial leaf fates. Encodes a member of HD-Zip family which contains homeodomain-leucine zipper domains and domain similar to a mammalian sterol binding domain. Has overlapping functions with PHAVOLUTA, REVOLUTA and CORONA. |
| Gma.16700.1.S1_at | 0.066588 | 5.225236 | Glyma20g16920.1 | AT5G61590.1 | encodes a member of the ERF (ethylene response factor) subfamily B-3 of ERF/AP2 transcription factor family. The protein contains one AP2 domain. There are 18 members in this subfamily including ATERF-1, ATERF-2, AND ATERF-5. |
| GmaAffx.15955.1.S1_ | 0.053928 | 5.082179 | Glyma11g02060.1 | AT2G46410.1 | Nuclear-localized R3-type MYB transcription factor. Positive regulator of hair-cell differentiation. Preferentially transcribed in hairless cells. Moves from atrichoblasts into trichoblast via plasmodesmata in a tissue-specific mode. N-terminus and part of the Myb domain are required for this movement, with W76 playing a crucial role. Capability to increase the size-exclusion limit of plasmodesmata. Regulated by WEREWOLF. |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.64710.1.S1_ | 0.072917 | 4.64099 | Glyma02g07760.1 | AT5G18270.1 | | ANAC087; FUNCTIONS IN: transcription factor activity; INVOLVED IN: multicellular organismal development, regulation of transcription; LOCATED IN: cellular_component unknown; EXPRESSED IN: 12 plant structures; EXPRESSED DURING: LP.06 six leaves visible, 4 anthesis, LP.04 four leaves visible, C globular stage, petal differentiation and expansion stage; CONTAINS InterPro DOMAIN/s: No apical meristem (NAM) protein (InterPro:IPR003441); BEST Arabidopsis thaliana protein match is: anac046 (Arabidopsis NAC domain containing protein 46); transcription factor (TAIR:AT3G04060.1); Has 1640 Blast hits to 1638 proteins in 54 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 1640; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| Gma.4526.1.S1_at | 0.092507 | 4.137146 | Glyma10g05210.1 | AT2G37430.1 | | zinc finger (C2H2 type) family protein (ZAT11); FUNCTIONS IN: transcription factor activity, zinc ion binding, nucleic acid binding; INVOLVED IN: response to chitin, regulation of transcription; LOCATED IN: intracellular; EXPRESSED IN: cotyledon, root, petiole; CONTAINS InterPro DOMAIN/s: Zinc finger, C2H2-like (InterPro:IPR015880), Zinc finger, C2H2-type (InterPro:IPR007087); BEST Arabidopsis thaliana protein match is: zinc finger (C2H2 type) family protein (TAIR:AT3G53600.1); Has 4550 Blast hits to 3560 proteins in 108 species: Archae - 0; Bacteria - 0; Metazoa - 3995; Fungi - 10; Plants - 464; Viruses - 0; Other Eukaryotes - 81 (source: NCBI BLink). |
| GmaAffx.42543.1.A1_ | 0.059042 | 3.191799 | Glyma19g08510.1 | AT1G25580.1 | | Encodes suppressor of gamma response 1 (SOG1), a putative transcription factor governing multiple responses to DNA damage. |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.17786.1.S1_at | 0.087286 | 2.886904 | No Soybean Matc | AT4G35580.2 | | NAC transcription factor-like 9 (NTL9); FUNCTIONS IN: transcription factor activity; INVOLVED IN: regulation of transcription; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: No apical meristem (NAM) protein (InterPro:IPR003441); BEST Arabidopsis thaliana protein match is: no apical meristem (NAM) family protein (TAIR:AT1G33060.2); Has 1627 Blast hits to 1625 proteins in 54 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 1627; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| Gma.16807.1.S1_at | 0.04032 | 2.684973 | Glyma09g14880.1 | AT1G78600.1 | | LIGHT-REGULATED ZINC FINGER PROTEIN 1 (LZF1); FUNCTIONS IN: transcription factor activity, zinc ion binding; INVOLVED IN: chlorophyll biosynthetic process, chloroplast organization, anthocyanin biosynthetic process, regulation of photomorphogenesis, regulation of transcription; LOCATED IN: nuclear speck; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Zinc finger, B-box (InterPro:IPR000315); BEST Arabidopsis thaliana protein match is: STO (SALT TOLERANCE); DNA binding / protein binding / transcription factor/ zinc ion binding (TAIR:AT1G06040.2); Has 1252 Blast hits to 916 proteins in 83 species: Archae - 0; Bacteria - 0; Metazoa - 17; Fungi - 0; Plants - 1143; Viruses - 0; Other Eukaryotes - 92 (source: NCBI BLink). |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.7454.1.S1_a_at | 0.065515 | 2.608343 | Glyma13g43800.1 | AT1G04250.1 | | Transcription regulator acting as repressor of auxin-inducible gene expression. Auxin-inducible AUX/IAA gene. Short-lived nuclear protein with four conserved domains. Domain III has homology to beta alpha alpha dimerization and DNA binding domains. Involved in auxin signaling. Auxin induces the degradation of the protein in a dosage-dependent manner in a process mediated by AtRac1. Auxin induced the relocalization of the protein within the nucleus from a diffused nucleoplasmic pattern to a discrete particulated pattern named nuclear protein bodies or NPB in a process also mediated by Rac1. Colocalizes with SCF, CSN and 26S proteasome components. |
| Gma.13973.1.A1_at | 0.072849 | 2.595982 | Glyma05g02550.1 | AT1G09540.1 | | Encodes putative transcription factor. Mutants lack of mucilage extrusion from the seeds during imbibition. Reduced quantities of mucilage are deposited during the development of the seed coat epidermis in myb61 mutants. Expressed in guard cells,loss of function mutations show an increase in stomatal pore opening suggesting a role in ABA independent regulation of stomatal pore size. |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.5448.1.S1_a | 0.086296 | 2.520167 | Glyma16g04720.1 | AT5G18270.1 | ANAC087; FUNCTIONS IN: transcription factor activity; INVOLVED IN: multicellular organismal development, regulation of transcription; LOCATED IN: cellular_component unknown; EXPRESSED IN: 12 plant structures; EXPRESSED DURING: LP.06 six leaves visible, 4 anthesis, LP.04 four leaves visible, C globular stage, petal differentiation and expansion stage; CONTAINS InterPro DOMAIN/s: No apical meristem (NAM) protein (InterPro:IPR003441); BEST Arabidopsis thaliana protein match is: anac046 (Arabidopsis NAC domain containing protein 46); transcription factor (TAIR:AT3G04060.1); Has 1640 Blast hits to 1638 proteins in 54 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 1640; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| Gma.8224.1.A1_at | 0.062003 | 2.43185 | Glyma17g15610.1 | AT4G17570.1 | Encodes a member of the GATA factor family of zinc finger transcription factors. |
| Gma.268.1.S1_at | 0.06006 | 2.279291 | Glyma17g35930.1 | AT1G32510.1 | Arabidopsis NAC domain containing protein 11 (anac011); FUNCTIONS IN: transcription factor activity; INVOLVED IN: regulation of transcription; LOCATED IN: cellular_component unknown; CONTAINS InterPro DOMAIN/s: No apical meristem (NAM) protein (InterPro:IPR003441); BEST Arabidopsis thaliana protein match is: anac071 (Arabidopsis NAC domain containing protein 71); transcription factor (TAIR:AT4G17980.1); Has 1546 Blast hits to 1544 proteins in 54 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 1546; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.4205.1.S1_at | 0.031716 | 2.242494 | Glyma06g23240.5 | AT5G23090.1 | | NUCLEAR FACTOR Y, SUBUNIT B13 (NF-YB13); FUNCTIONS IN: transcription factor activity; INVOLVED IN: regulation of transcription; LOCATED IN: nucleus; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Transcription factor CBF/NF-Y/archaeal histone (InterPro:IPR003958), Histone-fold (InterPro:IPR009072); BEST Arabidopsis thaliana protein match is: NF-YB12 (NUCLEAR FACTOR Y, SUBUNIT B12); DNA binding / transcription factor (TAIR:AT5G08190.1); Has 987 Blast hits to 987 proteins in 177 species: Archae - 0; Bacteria - 0; Metazoa - 384; Fungi - 235; Plants - 292; Viruses - 0; Other Eukaryotes - 76 (source: NCBI BLink). |
| GmaAffx.38321.1.S1_ | 0.056028 | 2.149361 | Glyma05g38380.1 | AT5G13180.1 | | Encodes a NAC domain transcription factor that interacts with VND7 and negatively regulates xylem vessel formation. |
| GmaAffx.65885.1.A1_ | 0.089272 | 1.95883 | Glyma01g41780.1 | AT2G45120.1 | | zinc finger (C2H2 type) family protein; FUNCTIONS IN: transcription factor activity, zinc ion binding, nucleic acid binding; INVOLVED IN: regulation of transcription; LOCATED IN: intracellular; CONTAINS InterPro DOMAIN/s: Zinc finger, C2H2-like (InterPro:IPR015880), Zinc finger, C2H2-type (InterPro:IPR007087); BEST Arabidopsis thaliana protein match is: zinc finger (C2H2 type) family protein (TAIR:AT3G60580.1); Has 2160 Blast hits to 2024 proteins in 149 species: Archae - 0; Bacteria - 8; Metazoa - 1372; Fungi - 66; Plants - 537; Viruses - 18; Other Eukaryotes - 159 (source: NCBI BLink). |
| GmaAffx.42667.3.S1_ | 0.051579 | 1.92235 | Glyma14g27290.1 | AT1G21450.1 | | Encodes a scarecrow-like protein (SCL1). Member of GRAS gene family. |
| GmaAffx.68318.1.S1_ | 0.051377 | 1.912293 | Glyma19g35180.4 | AT2G33310.3 | | Auxin induced gene, IAA13 (IAA13). |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.90028.1.S1_ | 0.089453 | 1.879673 | Glyma04g38560.1 | AT1G01720.1 | | Belongs to a large family of putative transcriptional activators with NAC domain. Transcript level increases in response to wounding and abscisic acid. ATAF1 attentuates ABA signaling and sythesis. Mutants are hyposensitive to ABA. |
| GmaAffx.79390.1.S1_ | 0.060214 | 1.867118 | Ambiguous Hit | AT1G08320.3 | | bZIP family transcription factor; FUNCTIONS IN: transcription factor activity, DNA binding; INVOLVED IN: regulation of transcription, DNA-dependent; LOCATED IN: nucleus; EXPRESSED IN: 9 plant structures; EXPRESSED DURING: 6 growth stages; CONTAINS InterPro DOMAIN/s: Basic-leucine zipper (bZIP) transcription factor (InterPro:IPR004827), bZIP transcription factor, bZIP-1 (InterPro:IPR011616); BEST Arabidopsis thaliana protein match is: bZIP family transcription factor (TAIR:AT5G06839.1); Has 758 Blast hits to 757 proteins in 88 species: Archae - 0; Bacteria - 0; Metazoa - 121; Fungi - 28; Plants - 577; Viruses - 0; Other Eukaryotes - 32 (source: NCBI BLink). |
| Gma.7043.1.S1_at | 0.06315 | 1.851152 | Glyma08g08220.1 | AT1G03970.1 | | encodes a basic leucine zipper G-box binding factor that can bind to G-box motifs only as heterodimers with GBF2 or GBF3. A single amino acid change can confer G-box binding as homodimers. |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.9475.1.A1_s | 0.039276 | 1.845075 | Glyma06g14290.1 | AT1G33060.1 | | ARABIDOPSIS NAC 014 (ANAC014); FUNCTIONS IN: transcription factor activity; INVOLVED IN: multicellular organismal development, regulation of transcription; LOCATED IN: cellular_component unknown; EXPRESSED IN: 14 plant structures; EXPRESSED DURING: 8 growth stages; CONTAINS InterPro DOMAIN/s: No apical meristem (NAM) protein (InterPro:IPR003441); BEST Arabidopsis thaliana protein match is: NTL9 (NAC transcription factor-like 9); transcription factor (TAIR:AT4G35580.1); Has 1631 Blast hits to 1628 proteins in 60 species: Archae - 0; Bacteria - 0; Metazoa - 4; Fungi - 0; Plants - 1621; Viruses - 5; Other Eukaryotes - 1 (source: NCBI BLink). |
| GmaAffx.40271.1.S1_ | 0.060127 | 1.837868 | Glyma06g12700.1 | AT1G21450.1 | | Encodes a scarecrow-like protein (SCL1). Member of GRAS gene family. |
| GmaAffx.65885.1.A1_ | 0.053551 | 1.773928 | Glyma01g41780.1 | AT2G45120.1 | | zinc finger (C2H2 type) family protein; FUNCTIONS IN: transcription factor activity, zinc ion binding, nucleic acid binding; INVOLVED IN: regulation of transcription; LOCATED IN: intracellular; CONTAINS InterPro DOMAIN/s: Zinc finger, C2H2-like (InterPro:IPR015880), Zinc finger, C2H2-type (InterPro:IPR007087); BEST Arabidopsis thaliana protein match is: zinc finger (C2H2 type) family protein (TAIR:AT3G60580.1); Has 2160 Blast hits to 2024 proteins in 149 species: Archae - 0; Bacteria - 8; Metazoa - 1372; Fungi - 66; Plants - 537; Viruses - 18; Other Eukaryotes - 159 (source: NCBI BLink). |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.7381.1.S1_at | 0.034734 | 1.669216 | Glyma20g33430.1 | AT3G10480.1 | | ARABIDOPSIS NAC DOMAIN CONTAINING PROTEIN 50 (ANAC050); FUNCTIONS IN: transcription factor activity; INVOLVED IN: multicellular organismal development, regulation of transcription; EXPRESSED IN: cultured cell; CONTAINS InterPro DOMAIN/s: No apical meristem (NAM) protein (InterPro:IPR003441); BEST Arabidopsis thaliana protein match is: anac052 (Arabidopsis NAC domain containing protein 52); transcription factor (TAIR:AT3G10490.2); Has 1719 Blast hits to 1708 proteins in 79 species: Archae - 0; Bacteria - 5; Metazoa - 32; Fungi - 8; Plants - 1608; Viruses - 0; Other Eukaryotes - 66 (source: NCBI BLink). |
| Gma.9082.1.S1_at | 0.071189 | 1.483862 | Glyma15g19980.1 | AT1G19220.1 | | Encodes an auxin response factor that contains the conserved VP1-B3 DNA-binding domain at its N-terminus and the Aux/IAA-like domains III and IV present in most ARFs at its C-terminus. The protein interacts with IAA1 (yeast two hybrid) and other auxin response elements such as ER7 and ER9 (yeast one hybrid). ARF19 protein can complement many aspects of the arf7 mutant phenotype and , together with ARF7, is involved in the response to ethylene. In the arf7 arf19 double mutant, several auxin-responsive genes (e.g. IAA5, LBD16, LBD29 and LBD33) are no longer upregulated by auxin. |

| | | | | | |
|---|---|---|---|---|---|
| Gma.1489.1.S1_at | 0.047594 | 1.442894 | No Soybean Matc | AT4G38900.3 | bZIP protein; FUNCTIONS IN: transcription factor activity, DNA binding; INVOLVED IN: regulation of transcription, DNA-dependent; LOCATED IN: nucleus; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Basic-leucine zipper (bZIP) transcription factor (InterPro:IPR004827), Basic leucine zipper (InterPro:IPR011700); BEST Arabidopsis thaliana protein match is: bZIP family transcription factor (TAIR:AT2G21230.1); Has 3065 Blast hits to 2187 proteins in 243 species: Archae - 8; Bacteria - 190; Metazoa - 623; Fungi - 196; Plants - 813; Viruses - 1; Other Eukaryotes - 1234 (source: NCBI BLink). |
| GmaAffx.9286.1.S1_s | 0.080963 | 1.379122 | Glyma11g14490.2 | AT3G46640.2 | Encodes a myb family transcription factor with a single Myb DNA-binding domain (type SHAQKYF) that is unique to plants and is essential for circadian rhythms, specifically for transcriptional regulation within the circadian clock. LUX is required for normal rhythmic expression of multiple clock outputs in both constant light and darkness. It is coregulated with TOC1 and seems to be repressed by CCA1 and LHY by direct binding of these proteins to the evening element in the LUX promoter. |
| Gma.3028.1.S1_at | 0.051714 | 1.362773 | Glyma06g17780.1 | AT3G48590.1 | Encodes a protein with similarity to a subunit of the CCAAT promoter motif binding complex of yeast.One of two members of this class (HAP5A) and expressed in vegetative and reproductive tissues. |
| GmaAffx.24129.1.S1_ | 0.063473 | 1.357012 | Glyma13g01930.1 | AT1G78080.1 | Encodes a member of the DREB subfamily A-6 of ERF/AP2 transcription factor family (RAP2.4). The protein contains one AP2 domain. Role in mediating light and ethylene signaling. |
| GmaAffx.8927.1.S1_a | 0.087219 | 1.334184 | Glyma06g01180.3 | AT4G39100.1 | Putative transcription factor containing a PHD finger and BAH motif, required for normal development |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.12424.1.S1_s_a | 0.096834 | 1.227917 | Glyma16g05840.1 | AT3G27010.1 | | Belongs to a TCP protein transcription factor family. Members of this family contain a predicted basic-helix-loop-helix domain involved in DNA binding. Related to rice PCF1 and PCF2 genes. Binds to the GCCCR element of CYCB1;1. Involved in regulation of expression of cell cycle control and ribosomal protein genes. |
| GmaAffx.39527.1.S1_ | 0.031716 | -8.936705 | Glyma11g09370.2 | AT2G17410.1 | | ARID/BRIGHT DNA-binding domain-containing protein; FUNCTIONS IN: transcription factor activity; INVOLVED IN: regulation of transcription; LOCATED IN: intracellular; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Heat shock protein Hsp20 (InterPro:IPR002068), HSP20-like chaperone (InterPro:IPR008978), AT-rich interaction region (InterPro:IPR001606); BEST Arabidopsis thaliana protein match is: ARID/BRIGHT DNA-binding domain-containing protein (TAIR:AT1G76510.2); Has 4562 Blast hits to 3323 proteins in 373 species: Archae - 17; Bacteria - 590; Metazoa - 1793; Fungi - 581; Plants - 192; Viruses - 45; Other Eukaryotes - 1344 (source: NCBI BLink). |
| GmaAffx.4921.1.S1_a | 0.075851 | -4.364399 | Glyma07g05950.1 | AT3G61850.4 | | Zinc finger transcription factor of the Dof family involved in the control of seed germination. |
| GmaAffx.70008.1.S1_ | 0.031716 | -3.888838 | Ambiguous Hit | AT1G66230.1 | | Encodes a putative transcription factor (MYB20). |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.8205.1.S1_a | 0.068696 | -3.666265 | Glyma12g34490.1 | AT3G20640.1 | | ethylene-responsive protein -related; FUNCTIONS IN: transcription factor activity; INVOLVED IN: regulation of transcription; LOCATED IN: nucleus, chloroplast; CONTAINS InterPro DOMAIN/s: Basic helix-loop-helix dimerisation region bHLH (InterPro:IPR001092), Helix-loop-helix DNA-binding (InterPro:IPR011598); BEST Arabidopsis thaliana protein match is: basic helix-loop-helix (bHLH) family protein (TAIR:AT1G61660.1); Has 2718 Blast hits to 1146 proteins in 107 species: Archae - 0; Bacteria - 21; Metazoa - 331; Fungi - 82; Plants - 843; Viruses - 0; Other Eukaryotes - 1441 (source: NCBI BLink). |
| GmaAffx.29616.1.A1_ | 0.068696 | -3.370755 | Glyma11g04440.2 | AT4G18020.4 | | Encodes pseudo-response regulator 2 (APRR2). |
| Gma.8383.1.S1_at | 0.04164 | -3.327861 | Glyma08g28220.1 | AT3G30530.1 | | ARABIDOPSIS THALIANA BASIC LEUCINE-ZIPPER 42 (ATBZIP42); FUNCTIONS IN: transcription factor activity, protein heterodimerization activity, protein homodimerization activity, DNA binding; INVOLVED IN: regulation of transcription, DNA-dependent; LOCATED IN: nucleus; EXPRESSED IN: inflorescence meristem, flower; EXPRESSED DURING: petal differentiation and expansion stage; CONTAINS InterPro DOMAIN/s: Basic-leucine zipper (bZIP) transcription factor (InterPro:IPR004827), bZIP transcription factor, bZIP-1 (InterPro:IPR011616); BEST Arabidopsis thaliana protein match is: AtbZIP43 (Arabidopsis thaliana basic leucine-zipper 43); DNA binding / transcription factor (TAIR:AT5G38800.1); Has 1055 Blast hits to 1055 proteins in 117 species: Archae - 0; Bacteria - 2; Metazoa - 126; Fungi - 18; Plants - 863; Viruses - 0; Other Eukaryotes - 46 (source: NCBI BLink). |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.3066.1.S1_at | 0.058702 | -3.042207 | Glyma19g40070.1 | AT2G47520.1 | | encodes a member of the ERF (ethylene response factor) subfamily B-2 of ERF/AP2 transcription factor family. The protein contains one AP2 domain. There are 5 members in this subfamily including RAP2.2 AND RAP2.12. |
| Gma.11276.1.S1_at | 0.051715 | -2.902341 | Glyma14g37050.1 | AT5G58900.1 | | myb family transcription factor; FUNCTIONS IN: transcription factor activity, DNA binding; INVOLVED IN: regulation of transcription; EXPRESSED IN: 20 plant structures; EXPRESSED DURING: 11 growth stages; CONTAINS InterPro DOMAIN/s: Molecular chaperone, heat shock protein, Hsp40, DnaJ (InterPro:IPR015609), SANT, eukarya (InterPro:IPR017884), Myb-like DNA-binding region, SHAQKYF class (InterPro:IPR006447), SANT, DNA-binding (InterPro:IPR001005), Homeodomain-like (InterPro:IPR009057), Myb, DNA-binding (InterPro:IPR014778), Homeodomain-related (InterPro:IPR012287), Myb-type HTH DNA-binding domain (InterPro:IPR017930); BEST Arabidopsis thaliana protein match is: myb family transcription factor (TAIR:AT2G38090.1); Has 1021 Blast hits to 1019 proteins in 106 species: Archae - 0; Bacteria - 0; Metazoa - 122; Fungi - 2; Plants - 822; Viruses - 0; Other Eukaryotes - 75 (source: NCBI BLink). |
| GmaAffx.68489.1.S1_ | 0.084461 | -2.88448 | Glyma08g17140.1 | AT1G79580.3 | | |
| GmaAffx.83173.1.S1_ | 0.06315 | -2.60627 | Glyma14g09310.2 | AT1G46480.1 | | Encodes a WUSCHEL-related homeobox gene family member with 65 amino acids in its homeodomain. Proteins in this family contain a sequence of eight residues (TLPLFPMH) downstream of the homeodomain called the WUS box. This protein also contains an acidic domain approximately 10 residues upstream of the WUS box. |

| | | | | | |
|---|---|---|---|---|---|
| Gma.12503.1.S1_at | 0.066512 | -2.487157 | Glyma01g02390.1 | AT5G65640.1 | beta HLH protein 93 (bHLH093); FUNCTIONS IN: transcription factor activity, DNA binding; INVOLVED IN: regulation of transcription; LOCATED IN: nucleus; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Basic helix-loop-helix dimerisation region bHLH (InterPro:IPR001092), Helix-loop-helix DNA-binding (InterPro:IPR011598); BEST Arabidopsis thaliana protein match is: basic helix-loop-helix (bHLH) family protein (TAIR:AT5G10570.1); Has 1575 Blast hits to 1575 proteins in 121 species: Archae - 0; Bacteria - 0; Metazoa - 74; Fungi - 13; Plants - 1485; Viruses - 0; Other Eukaryotes - 3 (source: NCBI BLink). |
| GmaAffx.26685.1.A1_ | 0.096225 | -2.234272 | Glyma19g36200.1 | AT2G28550.1 | RELATED TO AP2.7 (RAP2.7); FUNCTIONS IN: transcription factor activity, DNA binding; INVOLVED IN: organ morphogenesis, regulation of transcription, DNA-dependent, vegetative to reproductive phase transition; LOCATED IN: nucleus; EXPRESSED IN: 15 plant structures; EXPRESSED DURING: 9 growth stages; CONTAINS InterPro DOMAIN/s: DNA-binding, integrase-type (InterPro:IPR016177), Pathogenesis-related transcriptional factor and ERF, DNA-binding (InterPro:IPR001471); BEST Arabidopsis thaliana protein match is: TOE2; DNA binding / transcription factor (TAIR:AT5G60120.1); Has 3106 Blast hits to 2890 proteins in 180 species: Archae - 0; Bacteria - 10; Metazoa - 0; Fungi - 0; Plants - 3041; Viruses - 7; Other Eukaryotes - 48 (source: NCBI BLink). |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.5405.1.S1_at | 0.042977 | -2.225108 | Glyma19g36200.1 | AT2G28550.1 | | RELATED TO AP2.7 (RAP2.7); FUNCTIONS IN: transcription factor activity, DNA binding; INVOLVED IN: organ morphogenesis, regulation of transcription, DNA-dependent, vegetative to reproductive phase transition; LOCATED IN: nucleus; EXPRESSED IN: 15 plant structures; EXPRESSED DURING: 9 growth stages; CONTAINS InterPro DOMAIN/s: DNA-binding, integrase-type (InterPro:IPR016177), Pathogenesis-related transcriptional factor and ERF, DNA-binding (InterPro:IPR001471); BEST Arabidopsis thaliana protein match is: TOE2; DNA binding / transcription factor (TAIR:AT5G60120.1); Has 3106 Blast hits to 2890 proteins in 180 species: Archae - 0; Bacteria - 10; Metazoa - 0; Fungi - 0; Plants - 3041; Viruses - 7; Other Eukaryotes - 48 (source: NCBI BLink). |
| GmaAffx.9569.2.S1_a | 0.074269 | -2.206679 | Glyma13g20370.2 | AT4G30080.1 | | Involved in root cap cell differentiation. Gene expression is regulated by mir160.Located in the nucleus. |
| Gma.2134.1.A1_at | 0.047196 | -1.992072 | Glyma07g27820.2 | AT1G66140.1 | | Encodes a zinc finger protein containing only a single zinc finger. |
| GmaAffx.60862.1.A1 | 0.069248 | -1.804924 | Glyma07g31870.1 | AT1G28310.2 | | Dof-type zinc finger domain-containing protein; FUNCTIONS IN: transcription factor activity, DNA binding; INVOLVED IN: regulation of transcription; CONTAINS InterPro DOMAIN/s: Zinc finger, Dof-type (InterPro:IPR003851); BEST Arabidopsis thaliana protein match is: Dof-type zinc finger domain-containing protein (TAIR:AT5G02460.1); Has 945 Blast hits to 838 proteins in 78 species: Archae - 0; Bacteria - 20; Metazoa - 141; Fungi - 16; Plants - 623; Viruses - 0; Other Eukaryotes - 145 (source: NCBI BLink). |
| GmaAffx.24871.1.S1_ | 0.067425 | -1.785221 | Glyma07g36180.1 | AT2G47070.1 | | member of SPL gene family, encodes DNA binding proteins and putative transcription factors. All have the SBP-box, which encodes the SBP-domain, required for and sufficient for interaction with DNA. |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.1954.1.S1_at | 0.095497 | -1.603148 | Glyma16g02680.1 | AT1G01250.1 | | encodes a member of the DREB subfamily A-4 of ERF/AP2 transcription factor family. The protein contains one AP2 domain. There are 17 members in this subfamily including TINY. |
| Gma.15643.2.S1_at | 0.095497 | -1.579742 | Glyma01g00600.1 | AT5G56840.1 | | DNA-binding family protein; FUNCTIONS IN: transcription factor activity, DNA binding, nucleic acid binding; INVOLVED IN: regulation of transcription; CONTAINS InterPro DOMAIN/s: SANT, DNA-binding (InterPro:IPR001005), Homeodomain-like (InterPro:IPR009057), Myb, DNA-binding (InterPro:IPR014778), Zinc finger, CCHC-type (InterPro:IPR001878), Myb-type HTH DNA-binding domain (InterPro:IPR017930), Myb-like DNA-binding region, SHAQKYF class (InterPro:IPR006447); BEST Arabidopsis thaliana protein match is: myb family transcription factor (TAIR:AT3G16350.1); Has 519 Blast hits to 518 proteins in 49 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 473; Viruses - 0; Other Eukaryotes - 46 (source: NCBI BLink). |
| Gma.3507.1.S1_at | 0.074757 | -1.516657 | Glyma05g36410.1 | AT4G12750.1 | | sequence-specific DNA binding / transcription factor; FUNCTIONS IN: transcription factor activity, sequence-specific DNA binding; INVOLVED IN: regulation of transcription, DNA-dependent; LOCATED IN: nucleus; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: DDT superfamily (InterPro:IPR018501), Homeobox (InterPro:IPR001356), DDT family (InterPro:IPR004022), DDT subgroup (InterPro:IPR018500); BEST Arabidopsis thaliana protein match is: HB-1 (homeobox-1); transcription factor (TAIR:AT1G28420.1); Has 70 Blast hits to 64 proteins in 13 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 61; Viruses - 0; Other Eukaryotes - 9 (source: NCBI BLink). |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.2721.1.S1_s | 0.04164 | -1.480733 | No Soybean Matc | AT5G37020.1 | | Encodes a member of the auxin response factor family. Mediates auxin response via expression of auxin regulated genes. Acts redundantly with ARF6 to control stamen elongation and flower maturation. Expression of ARF8 is controlled by miR167. |

Table 10 Candidate genes with unknow functionality

| Affy ID | Q-value | Fold change | Soybean Gene | Arabidopsis homolog | Description |
|---|---|---|---|---|---|
| GmaAffx.58519.1.A1_at | 0.0585045 | 51.20405 | Glyma10g41130.1 | AT2G01300.1 | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; EXPRESSED IN: 11 plant structures; EXPRESSED DURING: 8 growth stages; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT1G15010.1); Has 45 Blast hits to 44 proteins in 7 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 45; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| GmaAffx.85033.1.S1_at | 0.0514544 | 14.0679 | Glyma10g01850.1 | AT5G38700.1 | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: mitochondrion; EXPRESSED IN: 10 plant structures; EXPRESSED DURING: LP.04 four leaves visible, LP.10 ten leaves visible, petal differentiation and expansion stage, LP.12 twelve leaves visible, LP.08 eight leaves visible; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT4G02170.1); Has 41 Blast hits to 41 proteins in 9 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 41; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| GmaAffx.90419.1.S1_at | 0.0598854 | 12.16055 | Glyma08g18230.2 | AT1G05340.1 | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: chloroplast; EXPRESSED IN: 14 plant structures; EXPRESSED DURING: 7 growth stages; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT2G32210.1); Has 122 Blast hits to 122 proteins in 22 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 16; Plants - 106; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.50725.1.S1_at | 0.0683924 | 11.79438 | Glyma05g33340.1 | AT5G07330.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: mitochondrion; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT1G63060.1); Has 18 Blast hits to 18 proteins in 7 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 18; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| GmaAffx.22115.1.S1_at | 0.0537629 | 10.45086 | Glyma14g13770.1 | AT5G35090.1 | | unknown protein; LOCATED IN: cellular_component unknown; EXPRESSED IN: 6 plant structures; EXPRESSED DURING: L mature pollen stage, M germinated pollen stage, 4 anthesis, petal differentiation and expansion stage; Has 42 Blast hits to 42 proteins in 4 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 42; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| Gma.1636.1.S1_s_at | 0.0317164 | 9.654402 | Glyma15g40760.1 | AT1G05340.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: chloroplast; EXPRESSED IN: 14 plant structures; EXPRESSED DURING: 7 growth stages; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT2G32210.1); Has 122 Blast hits to 122 proteins in 22 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 16; Plants - 106; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| GmaAffx.93383.1.S1_at | 0.0596963 | 9.373893 | Glyma08g22530.1 | AT1G15385.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; Has 4 Blast hits to 4 proteins in 2 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 4; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| GmaAffx.19824.1.S1_at | 0.0231573 | 9.133161 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| Gma.12326.1.S1_at | 0.0231573 | 8.934124 | Glyma13g40500.1 | NO BLASTP HIT (E<10-6) | | |

| | | | | | |
|---|---|---|---|---|---|
| Gma.12932.1.A1_at | 0.0312607 | 8.393175 | Glyma06g14440.1 | AT5G24680.1 | FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; CONTAINS InterPro DOMAIN/s: Peptidase C78, ubiquitin fold modifier-specific peptidase 1 and 2 (InterPro:IPR012462); Has 157 Blast hits to 157 proteins in 66 species: Archae - 0; Bacteria - 0; Metazoa - 86; Fungi - 45; Plants - 19; Viruses 0; Other Eukaryotes - 7 (source: NCBI BLink). |
| Gma.818.1.A1_at | 0.0953117 | 8.348683 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | |
| GmaAffx.86667.1.S1_at | 0.0796862 | 8.276215 | Glyma20g26790.1 | AT3G19550.1 | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: mitochondrion; EXPRESSED IN: 15 plant structures; EXPRESSED DURING: 9 growth stages; Has 16 Blast hits to 16 proteins in 4 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 16; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| GmaAffx.42864.1.S1_at | 0.07755 | 6.99284 | Glyma05g36510.1 | NO BLASTP HIT (E<10-6) | |
| GmaAffx.81651.1.S1_at | 0.0935144 | 6.897424 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | |
| Gma.8048.1.S1_at | 0.0475936 | 6.622849 | Glyma08g16050.1 | NO BLASTP HIT (E<10-6) | |
| GmaAffx.73388.1.A1_at | 0.0500445 | 6.433494 | Glyma03g30660.1 | AT2G48090.1 | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: chloroplast; EXPRESSED IN: male gametophyte, flower, root, carpel; EXPRESSED DURING: 4 anthesis, petal differentiation and expansion stage; Has 3 Blast hits to 3 proteins in 1 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 3; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.12031.2.S1_at | 0.0267886 | 6.233566 | Glyma02g04000.1 | AT1G70420.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 19 plant structures; EXPRESSED DURING: 12 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF1645 (InterPro:IPR012442); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT1G23710.1); Has 116 Blast hits to 116 proteins in 15 species: Archae - 0; Bacteria - 0; Metazoa - 9; Fungi - 0; Plants - 100; Viruses - 0; Other Eukaryotes - 7 (source: NCBI BLink). |
| Gma.1636.1.S1_at | 0.0231573 | 6.09952 | Glyma15g40760.1 | AT1G05340.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: chloroplast; EXPRESSED IN: 14 plant structures; EXPRESSED DURING: 7 growth stages; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT2G32210.1); Has 122 Blast hits to 122 proteins in 22 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 16; Plants - 106; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| GmaAffx.56797.1.A1_at | 0.0267886 | 6.011474 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| Gma.3586.1.S1_at | 0.0267886 | 5.910491 | Glyma20g26140.1 | AT2G01300.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; EXPRESSED IN: 11 plant structures; EXPRESSED DURING: 8 growth stages; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT1G15010.1); Has 45 Blast hits to 44 proteins in 7 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 45; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| Gma.12804.1.A1_at | 0.094428 | 5.87121 | Glyma08g22080.1 | NO BLASTP HIT (E<10-6) | | |
| Gma.1805.1.S1_at | 0.0471957 | 5.437036 | Glyma13g37000.1 | AT2G20562.1 | | |
| Gma.17894.1.S1_at | 0.0466862 | 5.390609 | Glyma08g16300.1 | NO BLASTP HIT (E<10-6) | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.38161.1.S1_at | 0.0925148 | 5.35878 | Glyma17g03470.1 | AT5G26910.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: mitochondrion; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT3G58650.1); Has 1070 Blast hits to 565 proteins in 119 species: Archae - 4; Bacteria - 320; Metazoa - 171; Fungi - 85; Plants - 92; Viruses - 0; Other Eukaryotes - 398 (source: NCBI BLink). |
| GmaAffx.1304.1.A1_at | 0.0426264 | 5.193883 | Glyma07g15290.1 | NO BLASTP HIT (E<10-6) | | |
| Gma.6108.1.S1_at | 0.026969 | 5.052116 | Glyma19g31190.1 | AT5G04080.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: plasma membrane; EXPRESSED IN: 16 plant structures; EXPRESSED DURING: 9 growth stages; Has 77 Blast hits to 77 proteins in 10 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 4; Plants - 73; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| GmaAffx.77799.1.A1_at | 0.0341303 | 4.801514 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| Gma.12031.2.S1_x_at | 0.0623842 | 4.779448 | Glyma02g04000.1 | AT1G70420.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 19 plant structures; EXPRESSED DURING: 12 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF1645 (InterPro:IPR012442); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT1G23710.1); Has 116 Blast hits to 116 proteins in 15 species: Archae - 0; Bacteria - 0; Metazoa - 9; Fungi - 0; Plants - 100; Viruses - 0; Other Eukaryotes - 7 (source: NCBI BLink). |
| GmaAffx.69557.1.S1_at | 0.0655151 | 4.430392 | Glyma03g02790.1 | NO BLASTP HIT (E<10-6) | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.8034.2.S1_at | 0.07755 | 4.323836 | Glyma04g34330.1 | AT1G74730.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: chloroplast thylakoid membrane, chloroplast; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF1118 (InterPro:IPR009500); Has 41 Blast hits to 41 proteins in 10 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 40; Viruses - 0; Other Eukaryotes - 1 (source: NCBI BLink). |
| Gma.10959.2.S1_at | 0.0163886 | 4.270438 | Glyma09g41560.1 | AT5G01750.2 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: chloroplast; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF567 (InterPro:IPR007612); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT3G11740.1); Has 229 Blast hits to 229 proteins in 33 species: Archae - 0; Bacteria - 18; Metazoa - 0; Fungi - 21; Plants - 190; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| GmaAffx.67214.1.S1_at | 0.0695792 | 4.260978 | Glyma10g07930.1 | AT1G28510.1 | | FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Optic atrophy 3-like (InterPro:IPR010754); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT3G58150.1); Has 279 Blast hits to 279 proteins in 116 species: Archae - 0; Bacteria - 0; Metazoa - 133; Fungi - 74; Plants - 40; Viruses - 0; Other Eukaryotes - 32 (source: NCBI BLink). |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.30792.1.A1_at | 0.0623842 | 4.248843 | Glyma03g38420.1 | AT3G62070.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT2G46940.1); Has 73 Blast hits to 72 proteins in 6 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 73; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| Gma.17839.1.S1_at | 0.0267886 | 4.182499 | Glyma13g32930.1 | AT5G64400.1 | | FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: CHCH (InterPro:IPR010625); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT5G09570.1); Has 218 Blast hits to 216 proteins in 80 species: Archae - 0; Bacteria - 10; Metazoa - 66; Fungi - 60; Plants - 47; Viruses - 0; Other Eukaryotes - 35 (source: NCBI BLink). |
| Gma.1735.1.A1_at | 0.0416405 | 4.112048 | Glyma19g31470.1 | NO BLASTP HIT (E<10-6) | | |
| GmaAffx.79482.1.A1_at | 0.0580906 | 4.083303 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| Gma.16004.1.S1_at | 0.0725836 | 4.076255 | Glyma01g39890.1 | AT5G67370.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF1230 (InterPro:IPR009631); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT5G11840.1); Has 264 Blast hits to 264 proteins in 69 species: Archae - 0; Bacteria - 98; Metazoa - 0; Fungi - 0; Plants - 35; Viruses - 0; Other Eukaryotes - 131 (source: NCBI BLink). |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | | | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: chloroplast; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF567 (InterPro:IPR007612); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT3G11740.1); Has 229 Blast hits to 229 proteins in 33 species: Archae - 0; Bacteria - 18; Metazoa - 0; Fungi - 21; Plants - 190; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| Gma.10959.1.S1_a_at | 0.0315416 | 4.051475 | Glyma18g44120.1 | AT5G01750.2 | | |
| | | | | | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: chloroplast; EXPRESSED IN: 15 plant structures; EXPRESSED DURING: 6 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function Cys-rich (InterPro:IPR006461); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT3G10980.1); Has 81 Blast hits to 75 proteins in 14 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 7; Plants - 74; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| Gma.8984.1.S1_at | 0.0337667 | 4.048486 | Glyma18g02600.1 | AT5G05350.1 | | |
| | | | | | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 10 plant structures; EXPRESSED DURING: 4 anthesis, petal differentiation and expansion stage; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT1G67920.1); Has 7 Blast hits to 7 proteins in 1 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 7; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| GmaAffx.91479.1.S1_s_at | 0.0581114 | 4.045082 | Glyma08g05010.1 | AT1G24600.1 | | |
| GmaAffx.91266.1.S1_s_at | 0.0535408 | 4.01434 | Glyma06g14100.1 | NO BLASTP HIT (E<10-6) | | |
| GmaAffx.16106.1.S1_at | 0.0577019 | 3.985733 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.57085.1.S1_at | 0.0337518 | 3.967323 | Glyma13g22540.1 | AT5G25240.1 | | FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; BEST Arabidopsis thaliana protein match is: NHL repeat-containing protein (TAIR:AT5G14890.1); Has 23 Blast hits to 23 proteins in 7 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 23; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| GmaAffx.36442.1.S1_at | 0.0550054 | 3.954091 | No Soybean Match Iden | AT4G32750.1 | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 18 plant structures; EXPRESSED DURING: 9 growth stages; Has 17 Blast hits to 17 proteins in 5 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 17; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). | |
| Gma.2984.1.S1_at | 0.0594787 | 3.936736 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| GmaAffx.65259.1.A1_at | 0.0964718 | 3.913533 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| GmaAffx.86456.1.S1_at | 0.0781102 | 3.815187 | Glyma13g12460.1 | NO BLASTP HIT (E<10-6) | | |
| Gma.9117.1.S1_at | 0.0862961 | 3.808232 | Glyma01g37010.1 | AT2G17850.1 | FUNCTIONS IN: molecular_function unknown; INVOLVED IN: aging; LOCATED IN: endomembrane system; EXPRESSED IN: 9 plant structures; EXPRESSED DURING: 9 growth stages; CONTAINS InterPro DOMAIN/s: Rhodanese-like (InterPro:IPR001763); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT5G66170.2); Has 1960 Blast hits to 1954 proteins in 396 species: Archae - 22; Bacteria - 1018; Metazoa - 11; Fungi - 7; Plants - 116; Viruses - 0; Other Eukaryotes - 786 (source: NCBI BLink). | |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.12054.2.A1_at | 0.0560276 | 3.782836 | Glyma13g43970.1 | AT5G11420.1 | | FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cell wall, plant-type cell wall; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF642 (InterPro:IPR006946), Galactose-binding like (InterPro:IPR008979); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT5G25460.1); Has 185 Blast hits to 157 proteins in 11 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 185; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| GmaAffx.22164.1.S1_at | 0.0613482 | 3.759378 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| Gma.8628.1.S1_at | 0.0267886 | 3.724734 | Glyma05g25920.1 | AT5G46295.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: stem, cotyledon, root, pollen tube, leaf; EXPRESSED DURING: LP.06 six leaves visible; Has 6 Blast hits to 6 proteins in 2 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 6; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| Gma.4524.1.A1_at | 0.0788633 | 3.689835 | Glyma11g03240.1 | AT3G60780.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; EXPRESSED IN: 7 plant structures; EXPRESSED DURING: L mature pollen stage, M germinated pollen stage, 4 anthesis, petal differentiation and expansion stage; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF1442 (InterPro:IPR009902); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT2G45360.1); Has 51 Blast hits to 51 proteins in 5 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 51; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.7254.1.S1_at | 0.0623842 | 3.682542 | Glyma09g23170.1 | AT5G52900.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 20 plant structures; EXPRESSED DURING: 11 growth stages; Has 10 Blast hits to 10 proteins in 3 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 10; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| GmaAffx.38977.1.S1_s_at | 0.0575919 | 3.67765 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| GmaAffx.45454.1.S1_at | 0.0591846 | 3.655954 | Glyma09g33060.1 | AT4G35560.1 | | FUNCTIONS IN: molecular_function unknown; LOCATED IN: CUL4 RING ubiquitin ligase complex; EXPRESSED IN: 20 plant structures; EXPRESSED DURING: 7 growth stages; CONTAINS InterPro DOMAIN/s: WD40 repeat-like (InterPro:IPR011046), WD40 repeat (InterPro:IPR001680), WD40/YVTN repeat-like (InterPro:IPR015943), Lethal giant larvae (Lgl) like, C-terminal (InterPro:IPR013905); BEST Arabidopsis thaliana protein match is: transducin family protein / WD-40 repeat family protein (TAIR:AT5G05570.1); Has 562 Blast hits to 546 proteins in 129 species: Archae - 0; Bacteria - 20; Metazoa - 305; Fungi - 89; Plants - 43; Viruses - 0; Other Eukaryotes - 105 (source: NCBI BLink). |
| GmaAffx.64502.1.S1_at | 0.0714804 | 3.655448 | Glyma13g02220.1 | NO BLASTP HIT (E<10-6) | | |
| GmaAffx.51710.1.A1_at | 0.0962969 | 3.56073 | Glyma19g09060.1 | NO BLASTP HIT (E<10-6) | | |
| GmaAffx.91973.1.S1_s_at | 0.0196047 | 3.538068 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.82499.1.S1_at | 0.054583 | 3.524762 | Glyma08g27620.1 | AT3G29970.1 | germination protein-related; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: endomembrane system; EXPRESSED IN: 8 plant structures; EXPRESSED DURING: F mature embryo stage, 4 leaf senescence stage, petal differentiation and expansion stage, D bilateral stage, E expanded cotyledon stage; CONTAINS InterPro DOMAIN/s: B12D (InterPro:IPR010530); BEST Arabidopsis thaliana protein match is: senescence-associated protein, putative (TAIR:AT3G48140.1); Has 77 Blast hits to 77 proteins in 15 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 77; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| GmaAffx.53187.1.S1_at | 0.0581114 | 3.490645 | Glyma19g37200.1 | NO BLASTP HIT (E<10-6) | |
| Gma.17174.1.S1_s_at | 0.0581114 | 3.477986 | Glyma02g18260.4 | AT4G39235.1 | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 15 plant structures; EXPRESSED DURING: 10 growth stages; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT3G05570.1); Has 39 Blast hits to 39 proteins in 7 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 39; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| GmaAffx.15182.1.A1_at | 0.0337518 | 3.44263 | Glyma17g20850.1 | NO BLASTP HIT (E<10-6) | |
| GmaAffx.11400.1.A1_at | 0.097011 | 3.418911 | Glyma12g24980.1 | NO BLASTP HIT (E<10-6) | |
| GmaAffx.52493.1.A1_at | 0.043378 | 3.386056 | Glyma03g25650.3 | NO BLASTP HIT (E<10-6) | |
| GmaAffx.57166.1.S1_at | 0.0392765 | 3.349184 | Glyma02g06260.1 | NO BLASTP HIT (E<10-6) | |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.66991.1.S1_at | 0.0684677 | 3.312133 | Glyma17g09610.1 | AT5G62960.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 13 growth stages; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT1G10660.4); Has 89 Blast hits to 88 proteins in 12 species: Archae - 0; Bacteria - 0; Metazoa - 8; Fungi - 0; Plants - 79; Viruses - 0; Other Eukaryotes - 2 (source: NCBI BLink). |
| Gma.17894.1.S1_s_at | 0.0196047 | 3.276045 | Glyma08g16300.1 | NO BLASTP HIT (E<10-6) | | |
| Gma.11672.1.S1_at | 0.0673275 | 3.230376 | Glyma15g01100.2 | AT1G52565.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: endomembrane system; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT3G15760.1); Has 21 Blast hits to 21 proteins in 6 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 21; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| GmaAffx.88213.1.S1_at | 0.0369675 | 3.192731 | Glyma13g03290.1 | AT1G21830.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; Has 31 Blast hits to 31 proteins in 7 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 31; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| GmaAffx.44978.1.S1_at | 0.0831792 | 3.181896 | Glyma08g00930.1 | NO BLASTP HIT (E<10-6) | | |
| GmaAffx.79862.1.S1_at | 0.0872493 | 3.169843 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.5781.3.S1_at | 0.0709225 | 3.097125 | No Soybean Match Iden | AT5G42300.1 | | UBIQUITIN-LIKE PROTEIN 5 (UBL5); FUNCTIONS IN: molecular_function unknown; INVOLVED IN: protein modification process; LOCATED IN: cellular_component unknown; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Ubiquitin (InterPro:IPR000626); BEST Arabidopsis thaliana protein match is: ubiquitin family protein (TAIR:AT3G45180.1); Has 300 Blast hits to 300 proteins in 135 species: Archae - 0; Bacteria - 0; Metazoa - 142; Fungi - 77; Plants - 44; Viruses - 0; Other Eukaryotes - 37 (source: NCBI BLink). |
| Gma.5902.1.S1_at | 0.0523348 | 3.088416 | Glyma15g04590.1 | NO BLASTP HIT (E<10-6) | | |
| GmaAffx.28364.1.S1_s_at | 0.0771246 | 2.941546 | Glyma08g08280.1 | AT1G36990.1 | | unknown protein; LOCATED IN: chloroplast; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 14 growth stages; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT4G08510.1); Has 4597 Blast hits to 1552 proteins in 255 species: Archae - 0; Bacteria - 986; Metazoa - 1160; Fungi - 777; Plants - 63; Viruses - 27; Other Eukaryotes - 1584 (source: NCBI BLink). |
| GmaAffx.81704.1.S1_at | 0.0369749 | 2.940183 | Glyma07g04440.1 | NO BLASTP HIT (E<10-6) | | |
| GmaAffx.43569.1.S1_at | 0.0316856 | 2.93682 | Glyma09g31700.1 | NO BLASTP HIT (E<10-6) | | |
| GmaAffx.12909.1.A1_at | 0.0671737 | 2.897987 | Glyma11g12090.1 | AT4G34630.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 14 growth stages; Has 14 Blast hits to 14 proteins in 5 species: Archae - 0; Bacteria - 2; Metazoa - 0; Fungi - 0; Plants - 8; Viruses - 0; Other Eukaryotes - 4 (source: NCBI BLink). |
| GmaAffx.28300.1.S1_s_at | 0.0370705 | 2.820065 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| GmaAffx.78636.1.S1_at | 0.0435849 | 2.810413 | Glyma06g42950.1 | NO BLASTP HIT (E<10-6) | | |
| GmaAffx.91589.1.S1_at | 0.0872861 | 2.797138 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| GmaAffx.28991.1.A1_at | 0.0613297 | 2.789121 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.91617.1.S1_at | 0.04484 | 2.781563 | Glyma05g03920.1 | AT2G45360.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF1442 (InterPro:IPR009902); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT3G60780.1); Has 55 Blast hits to 55 proteins in 8 species: Archae - 0; Bacteria - 6; Metazoa - 0; Fungi - 0; Plants - 49; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| GmaAffx.51876.1.S1_at | 0.0649367 | 2.768428 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| Gma.12764.1.A1_at | 0.0534418 | 2.740142 | Glyma08g43040.1 | NO BLASTP HIT (E<10-6) | | |
| GmaAffx.35493.1.S1_at | 0.0446585 | 2.738542 | Glyma04g38620.2 | AT1G27990.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT5G52420.1); Has 43 Blast hits to 43 proteins in 8 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 43; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| GmaAffx.57621.1.S1_at | 0.0362282 | 2.73544 | Glyma08g45240.1 | AT1G07280.3 | | binding; FUNCTIONS IN: binding; INVOLVED IN: biological_process unknown; LOCATED IN: chloroplast; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Tetratricopeptide-like helical (InterPro:IPR011990), Tetratricopeptide region (InterPro:IPR013026); BEST Arabidopsis thaliana protein match is: binding (TAIR:AT2G29670.1); Has 381 Blast hits to 218 proteins in 33 species: Archae - 4; Bacteria - 98; Metazoa - 1; Fungi - 2; Plants - 229; Viruses - 0; Other Eukaryotes - 47 (source: NCBI BLink). |
| GmaAffx.69470.1.S1_at | 0.0523348 | 2.715649 | Glyma13g21070.1 | NO BLASTP HIT (E<10-6) | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.58552.1.A1_at | 0.0336887 | 2.711867 | Glyma17g15290.1 | AT5G47410.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 12 growth stages; Has 12 Blast hits to 12 proteins in 7 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 12; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| GmaAffx.35331.1.S1_at | 0.0604275 | 2.710883 | Glyma14g11240.1 | NO BLASTP HIT (E<10-6) | | |
| GmaAffx.38773.1.A1_at | 0.0815612 | 2.707635 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| GmaAffx.3758.1.S1_at | 0.0649631 | 2.690895 | Glyma08g08860.1 | AT5G46295.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: stem, cotyledon, root, pollen tube, leaf; EXPRESSED DURING: LP.06 six leaves visible; Has 6 Blast hits to 6 proteins in 2 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 6; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| GmaAffx.8400.1.A1_at | 0.0267886 | 2.688155 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| GmaAffx.58109.1.S1_at | 0.0594787 | 2.672582 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| GmaAffx.11685.1.S1_at | 0.0911988 | 2.664492 | Glyma15g24130.1 | AT5G25240.1 | | FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; BEST Arabidopsis thaliana protein match is: NHL repeat-containing protein (TAIR:AT5G14890.1); Has 23 Blast hits to 23 proteins in 7 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 23; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.90677.1.S1_s_at | 0.0267886 | 2.638895 | Glyma04g01490.1 | AT1G49310.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: endomembrane system; EXPRESSED IN: 9 plant structures; EXPRESSED DURING: LP.04 four leaves visible, petal differentiation and expansion stage; Has 4 Blast hits to 4 proteins in 2 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 4; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| GmaAffx.69502.1.S1_at | 0.034734 | 2.575874 | Glyma01g08650.1 | NO BLASTP HIT (E<10-6) | | |
| GmaAffx.70071.1.S1_at | 0.0661268 | 2.528808 | Glyma11g05250.1 | NO BLASTP HIT (E<10-6) | | |
| GmaAffx.761.1.S1_at | 0.0744068 | 2.52099 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| GmaAffx.5942.1.S1_at | 0.0369675 | 2.500572 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| GmaAffx.65743.1.S1_at | 0.0819009 | 2.475492 | Glyma02g09000.1 | AT1G74450.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF793 (InterPro:IPR008511); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT1G18740.1); Has 121 Blast hits to 121 proteins in 12 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 121; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| Gma.11430.1.S1_at | 0.07755 | 2.472386 | Glyma12g36700.1 | NO BLASTP HIT (E<10-6) | | |
| Gma.16644.1.S1_at | 0.0501682 | 2.451597 | Glyma12g04300.1 | AT4G34630.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 14 growth stages; Has 14 Blast hits to 14 proteins in 5 species: Archae - 0; Bacteria - 2; Metazoa - 0; Fungi - 0; Plants - 8; Viruses - 0; Other Eukaryotes - 4 (source: NCBI BLink). |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.26524.1.S1_at | 0.0919093 | 2.412209 | Ambiguous Hit | AT3G52105.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: endomembrane system; Has 16 Blast hits to 16 proteins in 7 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 16; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| GmaAffx.88407.1.S1_at | 0.0692475 | 2.409855 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| GmaAffx.24626.1.S1_at | 0.0684677 | 2.404401 | Glyma06g13830.2 | AT5G33406.1 | | protein dimerization; FUNCTIONS IN: protein dimerization activity; INVOLVED IN: biological_process unknown; LOCATED IN: mitochondrion; CONTAINS InterPro DOMAIN/s: HAT dimerisation (InterPro:IPR008906); BEST Arabidopsis thaliana protein match is: DNA binding / protein dimerization (TAIR:AT4G15020.2); Has 413 Blast hits to 388 proteins in 17 species: Archae - 0; Bacteria - 0; Metazoa - 6; Fungi - 1; Plants - 399; Viruses - 0; Other Eukaryotes - 7 (source: NCBI BLink). |
| GmaAffx.66392.1.S1_at | 0.0844613 | 2.335462 | Glyma20g30700.1 | AT1G74450.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF793 (InterPro:IPR008511); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT1G18740.1); Has 121 Blast hits to 121 proteins in 12 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 121; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| Gma.3350.1.A1_s_at | 0.0684677 | 2.329624 | No Soybean Match Iden | AT2G20562.1 | | unknown protein; LOCATED IN: endomembrane system; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT2G31090.1); Has 34 Blast hits to 34 proteins in 7 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 34; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| Gma.7411.1.A1_at | 0.0494177 | 2.297193 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| GmaAffx.86863.1.A1_at | 0.0499032 | 2.295185 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.133.1.S1_at | 0.0998527 | 2.283498 | Glyma12g00370.1 | AT1G22600.1 | | FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: endomembrane system; EXPRESSED IN: stem, sperm cell, embryo, hypocotyl, flower; EXPRESSED DURING: petal differentiation and expansion stage, D bilateral stage; BEST Arabidopsis thaliana protein match is: late embryogenesis abundant domain-containing protein / LEA domain-containing protein (TAIR:AT1G72100.1); Has 126 Blast hits to 123 proteins in 31 species: Archae - 0; Bacteria - 0; Metazoa - 63; Fungi - 8; Plants - 55; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| Gma.13298.1.S1_s_at | 0.0641208 | 2.276643 | Glyma02g47310.1 | AT3G13275.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: endomembrane system; Has 13 Blast hits to 13 proteins in 7 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 13; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| Gma.7579.1.S1_at | 0.0285653 | 2.273439 | Glyma17g15250.3 | AT5G47420.1 | | FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: endoplasmic reticulum; EXPRESSED IN: 20 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF124 (InterPro:IPR002838), Tryptophan RNA-binding attenuator protein-like (InterPro:IPR016031); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT4G17420.1); Has 568 Blast hits to 568 proteins in 252 species: Archae - 56; Bacteria - 410; Metazoa - 0; Fungi - 8; Plants - 37; Viruses - 0; Other Eukaryotes - 57 (source: NCBI BLink). |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.1640.1.S1_at | 0.0274337 | 2.265842 | Glyma11g08590.1 | AT2G17705.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: plasma membrane; EXPRESSED IN: cultured cell; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT4G08330.1); Has 74 Blast hits to 74 proteins in 21 species: Archae - 5; Bacteria - 14; Metazoa - 0; Fungi - 0; Plants - 45; Viruses - 0; Other Eukaryotes - 10 (source: NCBI BLink). |
| GmaAffx.88979.1.A1_s_at | 0.0336887 | 2.264395 | Glyma15g42800.1 | NO BLASTP HIT (E<10-6) | | |
| Gma.4759.1.A1_at | 0.0316856 | 2.195289 | Glyma20g30160.1 | NO BLASTP HIT (E<10-6) | | |
| Gma.12856.1.A1_at | 0.0373064 | 2.181459 | Glyma02g10130.2 | AT2G39650.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF506, plant (InterPro:IPR006502); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT4G14620.1); Has 215 Blast hits to 215 proteins in 16 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 213; Viruses - 0; Other Eukaryotes - 2 (source: NCBI BLink). |
| Gma.5258.1.S1_at | 0.0515786 | 2.181335 | Glyma02g38770.1 | NO BLASTP HIT (E<10-6) | | |
| GmaAffx.79541.1.S1_at | 0.0052604 | 2.148961 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| Gma.4551.1.S1_at | 0.0352773 | 2.138614 | Glyma04g06720.1 | AT4G32020.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: chloroplast; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT2G25250.1); Has 34 Blast hits to 34 proteins in 12 species: Archae - 0; Bacteria - 0; Metazoa - 1; Fungi - 5; Plants - 28; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.368.1.S1_at | 0.07755 | 2.131974 | Glyma08g15450.1 | AT2G22795.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT4G37820.1); Has 742033 Blast hits to 185230 proteins in 3414 species: Archae - 3102; Bacteria - 91104; Metazoa - 304818; Fungi - 76748; Plants - 32809; Viruses 4044; Other Eukaryotes - 229408 (source: NCBI BLink). |
| GmaAffx.53212.1.S1_at | 0.0600598 | 2.125661 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| GmaAffx.47778.1.A1_at | 0.097011 | 2.115492 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| GmaAffx.12463.2.S1_at | 0.0789281 | 2.114587 | Ambiguous Hit | NO BLASTX HIT (E<10-6) | | |
| GmaAffx.29900.1.S1_at | 0.0500445 | 2.113643 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| GmaAffx.58958.1.S1_at | 0.0490005 | 2.113442 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| Gma.17260.1.A1_at | 0.0317164 | 2.090934 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| Gma.10075.1.S1_at | 0.0800804 | 2.074594 | Glyma19g29030.2 | AT1G56660.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 16 plant structures; EXPRESSED DURING: 8 growth stages; Has 553849 Blast hits to 156892 proteins in 3110 species: Archae - 2183; Bacteria - 59979; Metazoa - 236712; Fungi - 59029; Plants - 23920; Viruses - 2741; Other Eukaryotes - 169285 (source: NCBI BLink). |
| Gma.1258.1.S1_at | 0.0868161 | 2.061724 | Glyma15g42780.1 | NO BLASTP HIT (E<10-6) | | |
| GmaAffx.78702.1.S1_at | 0.0731888 | 2.050446 | Glyma15g04440.1 | AT5G59350.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: endomembrane system; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; Has 21 Blast hits to 21 proteins in 6 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 21; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | | | | unknown protein; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 14 plant structures; EXPRESSED DURING: LP.06 six leaves visible, 4 anthesis, LP.04 four leaves visible, 4 leaf senescence stage, petal differentiation and expansion stage; Has 48 Blast hits to 48 proteins in 22 species: Archae - 0; Bacteria - 20; Metazoa - 11; Fungi - 2; Plants - 14; Viruses - 0; Other Eukaryotes - 1 (source: NCBI BLink). |
| GmaAffx.76696.1.A1_at | 0.0630429 | 2.042661 | Glyma18g32910.1 | AT4G13030.1 | | |
| Gma.16628.1.S1_at | 0.0595153 | 2.036643 | Glyma01g08580.1 | AT1G26300.1 | BSD domain-containing protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: BSD (InterPro:IPR005607); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT1G69030.1); Has 146 Blast hits to 143 proteins in 27 species: Archae - 2; Bacteria - 4; Metazoa - 9; Fungi - 4; Plants - 98; Viruses - 0; Other Eukaryotes - 29 (source: NCBI BLink). | |
| Gma.6305.1.S1_at | 0.0686459 | 2.02756 | Glyma06g07930.1 | AT4G31080.1 | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT2G24330.1); Has 251 Blast hits to 245 proteins in 93 species: Archae - 0; Bacteria - 0; Metazoa - 120; Fungi - 75; Plants - 53; Viruses - 0; Other Eukaryotes - 3 (source: NCBI BLink). | |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.66293.2.S1_at | 0.0661268 | 2.024468 | Glyma08g16020.2 | AT2G44500.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF246, plant (InterPro:IPR004348); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT3G07900.1); Has 429 Blast hits to 426 proteins in 22 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 428; Viruses - 0; Other Eukaryotes - 1 (source: NCBI BLink). |
| GmaAffx.80060.1.A1_at | 0.0370705 | 2.008276 | Glyma19g41160.1 | AT3G05940.1 | | unknown protein; INVOLVED IN: biological_process unknown; LOCATED IN: endomembrane system; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF300 (InterPro:IPR005178); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT5G26740.3); Has 669 Blast hits to 664 proteins in 143 species: Archae - 0; Bacteria - 0; Metazoa - 318; Fungi - 124; Plants - 129; Viruses - 0; Other Eukaryotes - 98 (source: NCBI BLink). |
| Gma.12400.1.S1_at | 0.0639467 | 1.998257 | Glyma05g08240.1 | AT1G52720.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT3G15630.1); Has 35 Blast hits to 35 proteins in 7 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 35; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.14679.1.S1_s_at | 0.0581114 | 1.995435 | Glyma17g06410.2 | AT4G29850.1 | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; CONTAINS InterPro DOMAIN/s: Uncharacterised protein family UPF0414 (InterPro:IPR008590); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT2G19350.1); Has 188 Blast hits to 188 proteins in 50 species: Archae - 0; Bacteria - 0; Metazoa - 141; Fungi - 0; Plants - 47; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| GmaAffx.1313.1.S1_at | 0.0796862 | 1.993343 | Glyma10g03030.2 | AT1G03290.1 | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT4G02880.1); Has 10945 Blast hits to 7852 proteins in 686 species: Archae - 147; Bacteria - 1290; Metazoa - 5004; Fungi - 733; Plants - 317; Viruses - 53; Other Eukaryotes - 3401 (source: NCBI BLink). |
| GmaAffx.67053.1.S1_s_at | 0.0623842 | 1.989739 | Glyma13g10640.1 | AT5G42050.1 | FUNCTIONS IN: molecular_function unknown; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Development and cell death domain (InterPro:IPR013989), Kelch related (InterPro:IPR013089); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT3G27090.1); Has 5084 Blast hits to 2870 proteins in 85 species: Archae - 0; Bacteria - 24; Metazoa - 270; Fungi - 58; Plants - 157; Viruses - 8; Other Eukaryotes - 4567 (source: NCBI BLink). |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.11276.1.S1_s_at | 0.0416405 | 1.981289 | No Soybean Match Iden | AT4G32480.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: cultured cell; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF506, plant (InterPro:IPR006502); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT2G20670.1); Has 203 Blast hits to 203 proteins in 16 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 201; Viruses - 0; Other Eukaryotes - 2 (source: NCBI BLink). |
| Gma.6758.1.S1_s_at | 0.0272696 | 1.979724 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| Gma.542.1.S1_at | 0.0416405 | 1.929364 | Glyma08g16310.1 | NO BLASTP HIT (E<10-6) | | |
| Gma.15091.1.S1_s_at | 0.0894528 | 1.916178 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| Gma.5709.1.S1_at | 0.0762235 | 1.904138 | Glyma10g42820.3 | AT4G35110.4 | | FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 15 plant structures; EXPRESSED DURING: 10 growth stages; CONTAINS InterPro DOMAIN/s: Phospholipase-like, arabidopsis (InterPro:IPR007942); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT2G16900.1); Has 412 Blast hits to 329 proteins in 86 species: Archae - 13; Bacteria - 24; Metazoa - 97; Fungi - 28; Plants - 128; Viruses - 35; Other Eukaryotes - 87 (source: NCBI BLink). |
| GmaAffx.70956.1.S1_at | 0.0390758 | 1.903991 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| Gma.8158.1.A1_s_at | 0.0585746 | 1.900888 | Glyma12g04760.2 | AT1G75180.3 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT1G19400.2); Has 70 Blast hits to 70 proteins in 8 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 70; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.62601.1.S1_at | 0.0824292 | 1.882058 | Glyma02g38390.1 | AT5G05190.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 19 plant structures; EXPRESSED DURING: 11 growth stages; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT3G56410.2); Has 4804 Blast hits to 1006 proteins in 190 species: Archae - 0; Bacteria - 237; Metazoa - 900; Fungi - 244; Plants - 143; Viruses - 5; Other Eukaryotes - 3275 (source: NCBI BLink). |
| GmaAffx.63468.1.A1_at | 0.0631501 | 1.879348 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| GmaAffx.16510.1.S1_at | 0.0510522 | 1.867403 | Glyma08g47320.1 | AT5G02440.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 17 plant structures; EXPRESSED DURING: 11 growth stages; Has 29 Blast hits to 29 proteins in 8 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 29; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| GmaAffx.22297.2.S1_at | 0.0822382 | 1.865181 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| Gma.17948.1.S1_at | 0.0910601 | 1.849741 | Glyma12g17960.1 | AT3G14870.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF641, plant (InterPro:IPR006943); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT1G53380.3); Has 291 Blast hits to 289 proteins in 49 species: Archae - 9; Bacteria - 4; Metazoa - 46; Fungi - 5; Plants - 180; Viruses - 0; Other Eukaryotes - 47 (source: NCBI BLink). |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.36562.1.S1_at | 0.0581114 | 1.842608 | Glyma07g34400.1 | AT2G01480.1 | | unknown protein; LOCATED IN: chloroplast; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF246, plant (InterPro:IPR004348); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT1G14970.1); Has 431 Blast hits to 423 proteins in 21 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 431; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| GmaAffx.13134.1.A1_s_at | 0.0633679 | 1.831228 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| Gma.1672.1.S1_at | 0.0649367 | 1.817138 | Glyma04g37640.3 | AT3G26890.5 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT5G41110.1); Has 276 Blast hits to 268 proteins in 89 species: Archae - 0; Bacteria - 34; Metazoa - 154; Fungi - 44; Plants - 39; Viruses - 0; Other Eukaryotes - 5 (source: NCBI BLink). |
| GmaAffx.60074.1.S1_at | 0.0471957 | 1.805464 | Glyma15g07490.2 | AT1G61240.4 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 19 plant structures; EXPRESSED DURING: 11 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF707 (InterPro:IPR007877); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT1G11170.1); Has 178 Blast hits to 178 proteins in 12 species: Archae - 0; Bacteria - 2; Metazoa - 0; Fungi - 0; Plants - 174; Viruses - 0; Other Eukaryotes - 2 (source: NCBI BLink). |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.24648.1.S1_at | 0.0851155 | 1.805204 | Glyma01g04560.1 | AT1G69980.1 | | unknown protein; LOCATED IN: endomembrane system; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; Has 15 Blast hits to 14 proteins in 8 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 15; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| Gma.4091.1.S1_a_at | 0.0711891 | 1.796359 | Ambiguous Hit | AT5G19080.1 | | zinc finger (C3HC4-type RING finger) family protein; |
| GmaAffx.71810.1.S1_at | 0.0637147 | 1.792049 | Glyma18g02810.1 | AT1G09870.1 | | histidine acid phosphatase family protein; FUNCTIONS IN: acid phosphatase activity; INVOLVED IN: biological_process unknown; LOCATED IN: plasma membrane; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Histidine acid phosphatase (InterPro:IPR000560), Histidine acid phosphatase, eukaryotic (InterPro:IPR016274); Has 574 Blast hits to 569 proteins in 162 species: Archae - 0; Bacteria - 75; Metazoa - 167; Fungi - 277; Plants - 29; Viruses - 0; Other Eukaryotes - 26 (source: NCBI BLink). |
| Gma.6517.1.S1_at | 0.0684677 | 1.790379 | Glyma09g24610.1 | AT4G25670.1 | | unknown protein; LOCATED IN: cellular_component unknown; EXPRESSED IN: male gametophyte, pollen tube; EXPRESSED DURING: M germinated pollen stage; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT4G25690.2); Has 46 Blast hits to 40 proteins in 10 species: Archae - 0; Bacteria - 0; Metazoa - 5; Fungi - 2; Plants - 38; Viruses - 0; Other Eukaryotes - 1 (source: NCBI BLink). |
| Gma.8324.1.S1_at | 0.0739194 | 1.789293 | Glyma01g02800.1 | AT5G65490.1 | | [illegible] IN: biological_process unknown; LOCATED IN: chloroplast; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: SGT1 (InterPro:IPR010770); Has 4480 Blast hits to 3390 proteins in 346 species: Archae - 8; Bacteria - 210; Metazoa - 1070; Fungi - 582; Plants - 124; Viruses - 61; Other Eukaryotes - 2425 (source: NCBI BLink). |
| Gma.11007.1.S1_a_at | 0.0377468 | 1.786946 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.26075.1.S1_at | 0.094028 | 1.76918 | Glyma15g03380.2 | AT1G06050.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF1336 (InterPro:IPR009769); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT5G10750.1); Has 180 Blast hits to 179 proteins in 16 species: Archae - 0; Bacteria - 0; Metazoa - 1; Fungi - 0; Plants - 169; Viruses - 0; Other Eukaryotes - 10 (source: NCBI BLink). |
| Gma.17907.1.A1_at | 0.0623842 | 1.751702 | Glyma10g02780.1 | NO BLASTP HIT (E<10-6) | | FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: chloroplast; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Isy1-like splicing (InterPro:IPR009360); Has 1075 Blast hits to 879 proteins in 176 species: Archae - 8; Bacteria - 11; Metazoa - 379; Fungi - 177; Plants - 27; Viruses - 9; Other Eukaryotes - 464 (source: NCBI BLink). |
| GmaAffx.55466.1.S1_at | 0.0810548 | 1.747642 | Glyma09g18830.1 | AT3G18790.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; Has 3355 Blast hits to 2831 proteins in 380 species: Archae - 39; Bacteria - 297; Metazoa - 1854; Fungi - 262; Plants - 105; Viruses - 3; Other Eukaryotes - 795 (source: NCBI BLink). |
| Gma.4454.2.A1_a_at | 0.0739694 | 1.73633 | Glyma19g05780.1 | AT5G15880.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; Has 18138 Blast hits to 9351 proteins in 574 species: Archae - 0; Bacteria - 666; Metazoa - 10236; Fungi - 1653; Plants - 1129; Viruses - 52; Other Eukaryotes - 4402 (source: NCBI BLink). |
| GmaAffx.60948.1.S1_at | 0.0452209 | 1.722493 | Glyma10g41300.2 | AT2G01100.3 | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.9862.1.A1_at | 0.0519648 | 1.721829 | Glyma03g24660.1 | AT3G60200.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT2G44600.1); Has 37 Blast hits to 37 proteins in 8 species: Archae - 0; Bacteria - 0; Metazoa - 8; Fungi - 0; Plants - 28; Viruses - 0; Other Eukaryotes - 1 (source: NCBI BLink). |
| GmaAffx.55942.1.S1_at | 0.0545932 | 1.671907 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| Gma.2498.1.A1_at | 0.0452209 | 1.66373 | Glyma14g04130.1 | AT2G26780.1 | | binding; FUNCTIONS IN: binding; INVOLVED IN: biological_process unknown; LOCATED IN: membrane; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: HEAT (InterPro:IPR000357), Armadillo-like helical (InterPro:IPR011989), Armadillo-type fold (InterPro:IPR016024); Has 325 Blast hits to 275 proteins in 96 species: Archae - 0; Bacteria - 0; Metazoa - 149; Fungi - 125; Plants - 34; Viruses - 0; Other Eukaryotes - 17 (source: NCBI BLink). |
| GmaAffx.36616.1.A1_at | 0.0695792 | 1.658424 | Glyma07g07460.1 | NO BLASTP HIT (E<10-6) | | |
| GmaAffx.87510.1.S1_at | 0.0811777 | 1.653756 | Glyma13g05780.1 | AT1G52530.1 | | FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 7 plant structures; EXPRESSED DURING: F mature embryo stage, petal differentiation and expansion stage, D bilateral stage, E expanded cotyledon stage; CONTAINS InterPro DOMAIN/s: Hus1-like protein (InterPro:IPR007150); Has 88 Blast hits to 87 proteins in 31 species: Archae - 0; Bacteria - 0; Metazoa - 51; Fungi - 6; Plants - 21; Viruses - 0; Other Eukaryotes - 10 (source: NCBI BLink). |
| Gma.17076.1.A1_at | 0.0623842 | 1.639968 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.6597.1.S1_at | 0.0461044 | 1.625423 | Glyma20g11560.1 | AT3G51980.1 | | binding; FUNCTIONS IN: binding; INVOLVED IN: biological_process unknown; LOCATED IN: endomembrane system; CONTAINS InterPro DOMAIN/s: Armadillo-like helical (InterPro:IPR011989), Armadillo-type fold (InterPro:IPR016024); BEST Arabidopsis thaliana protein match is: armadillo/beta-catenin repeat family protein (TAIR:AT3G09350.1); Has 470 Blast hits to 470 proteins in 146 species: Archae - 0; Bacteria - 0; Metazoa - 227; Fungi - 113; Plants - 80; Viruses - 0; Other Eukaryotes - 50 (source: NCBI BLink). |
| Gma.3857.1.S1_a_at | 0.094028 | 1.615723 | Ambiguous Hit | AT2G23790.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF607 (InterPro:IPR006769); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT4G36820.1); Has 267 Blast hits to 265 proteins in 87 species: Archae - 0; Bacteria - 0; Metazoa - 132; Fungi - 34; Plants - 71; Viruses - 0; Other Eukaryotes - 30 (source: NCBI BLink). |
| Gma.3899.1.A1_at | 0.0267886 | 1.613259 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| Gma.1358.1.S1_at | 0.0765822 | 1.595569 | Glyma01g02090.1 | AT4G17790.1 | | INVOLVED IN: biological_process unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: SNARE associated Golgi protein (InterPro:IPR015414); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT1G71940.1); Has 2258 Blast hits to 2258 proteins in 517 species: Archae - 4; Bacteria - 1129; Metazoa - 188; Fungi - 22; Plants - 104; Viruses - 0; Other Eukaryotes - 811 (source: NCBI BLink). |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.27358.1.S1_at | 0.0723382 | 1.593155 | Glyma12g14350.2 | AT2G31130.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 12 growth stages; Has 78 Blast hits to 77 proteins in 33 species: Archae - 0; Bacteria - 3; Metazoa - 17; Fungi - 2; Plants - 10; Viruses - 0; Other Eukaryotes - 46 (source: NCBI BLink). |
| Gma.2518.1.S1_a_at | 0.0576293 | 1.592283 | Glyma15g11070.1 | AT5G19070.1 | | LOCATED IN: endomembrane system; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: SNARE associated Golgi protein (InterPro:IPR015414); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT1G03260.1); Has 3029 Blast hits to 3029 proteins in 693 species: Archae - 10; Bacteria - 1724; Metazoa - 229; Fungi - 74; Plants - 161; Viruses - 0; Other Eukaryotes - 831 (source: NCBI BLink). |
| GmaAffx.35364.2.S1_at | 0.0549799 | 1.591864 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| Gma.11007.1.S1_x_at | 0.0500617 | 1.588446 | Glyma15g40790.2 | AT2G32190.2 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT2G32210.1); Has 40 Blast hits to 40 proteins in 4 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 40; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| Gma.10365.1.S1_at | 0.0451179 | 1.588337 | Glyma15g04910.1 | AT2G28570.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; EXPRESSED IN: inflorescence meristem, hypocotyl, flower, root; EXPRESSED DURING: petal differentiation and expansion stage; Has 2 Blast hits to 2 proteins in 1 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 2; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| GmaAffx.90854.1.S1_at | 0.0274337 | 1.572822 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.16586.1.S1_at | 0.0744068 | 1.561203 | Glyma13g10240.1 | AT3G54200.1 | | FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: anchored to plasma membrane, plasma membrane; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Harpin-induced 1 (InterPro:IPR010847); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT3G05975.1); Has 405 Blast hits to 404 proteins in 16 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 405; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| Gma.2426.2.S1_a_at | 0.0631501 | 1.559083 | Glyma04g09790.2 | AT5G49710.2 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT4G24590.1); Has 48 Blast hits to 48 proteins in 14 species: Archae - 0; Bacteria - 2; Metazoa - 3; Fungi - 0; Plants - 37; Viruses - 0; Other Eukaryotes - 6 (source: NCBI BLink). |
| GmaAffx.53010.1.A1_at | 0.0649367 | 1.558473 | Glyma12g02590.1 | AT2G41990.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 20 plant structures; EXPRESSED DURING: 13 growth stages; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT4G35170.1); Has 89 Blast hits to 88 proteins in 9 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 89; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| GmaAffx.38434.1.S1_at | 0.0711891 | 1.553149 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.88163.1.S1_at | 0.0948435 | 1.552408 | Glyma09g33850.1 | | AT4G17790.1 | INVOLVED IN: biological_process unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: SNARE associated Golgi protein (InterPro:IPR015414); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT1G71940.1); Has 2258 Blast hits to 2258 proteins in 517 species: Archae - 4; Bacteria - 1129; Metazoa - 188; Fungi - 22; Plants - 104; Viruses - 0; Other Eukaryotes - 811 (source: NCBI BLink). |
| Gma.5256.2.S1_at | 0.0765759 | 1.550893 | Glyma10g00250.1 | | AT3G01680.1 | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT3G01670.1); Has 49 Blast hits to 46 proteins in 10 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 49; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| Gma.9206.1.A1_at | 0.0892722 | 1.536729 | Glyma16g23180.1 | | NO BLASTP HIT (E<10-6) | |
| Gma.4299.1.S1_at | 0.0581114 | 1.534045 | Glyma09g12320.2 | | AT2G12400.1 | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: endomembrane system; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 13 growth stages; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT2G25270.1); Has 92 Blast hits to 88 proteins in 12 species: Archae - 0; Bacteria - 0; Metazoa - 2; Fungi - 0; Plants - 85; Viruses - 0; Other Eukaryotes - 5 (source: NCBI BLink). |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.2008.1.S1_at | 0.0595153 | 1.531376 | Glyma09g07030.1 | AT4G29850.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; CONTAINS InterPro DOMAIN/s: Uncharacterised protein family UPF0414 (InterPro:IPR008590); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT2G19350.1); Has 188 Blast hits to 188 proteins in 50 species: Archae - 0; Bacteria - 0; Metazoa - 141; Fungi - 0; Plants - 47; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| Gma.6511.1.S1_at | 0.0813937 | 1.524814 | Glyma01g36720.1 | AT2G17710.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; Has 19 Blast hits to 19 proteins in 9 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 19; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| GmaAffx.50095.2.A1_s_at | 0.0623842 | 1.498911 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| GmaAffx.13732.1.A1_at | 0.0970616 | 1.498887 | Glyma18g29560.1 | ATMG00310.1 | | hypothetical protein |
| Gma.11943.1.S1_a_at | 0.0616577 | 1.497235 | Glyma09g35820.1 | AT2G45010.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function Cys-rich (InterPro:IPR006461); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT5G51400.1); Has 322 Blast hits to 321 proteins in 35 species: Archae - 0; Bacteria - 0; Metazoa - 5; Fungi - 15; Plants - 299; Viruses - 0; Other Eukaryotes - 3 (source: NCBI BLink). |
| GmaAffx.27505.1.A1_s_at | 0.0771246 | 1.496182 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.16367.1.S1_at | 0.0797779 | 1.483426 | Glyma15g14760.1 | AT5G49540.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF786 (InterPro:IPR008504); Has 172 Blast hits to 172 proteins in 86 species: Archae - 0; Bacteria - 0; Metazoa - 102; Fungi - 27; Plants - 15; Viruses - 0; Other Eukaryotes - 28 (source: NCBI BLink). |
| GmaAffx.21063.1.A1_at | 0.0684677 | 1.479956 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| Gma.5699.1.S1_at | 0.0131636 | 1.479929 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| Gma.2698.1.S1_at | 0.0925065 | 1.448202 | Glyma20g25950.2 | AT2G01100.3 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; Has 18138 Blast hits to 9351 proteins in 574 species: Archae - 0; Bacteria - 666; Metazoa - 10236; Fungi - 1653; Plants - 1129; Viruses - 52; Other Eukaryotes - 4402 (source: NCBI BLink). |
| Gma.14515.3.S1_at | 0.0623842 | 1.445365 | Glyma05g13330.1 | AT1G31660.1 | | FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Bystin (InterPro:IPR007955); Has 370 Blast hits to 362 proteins in 156 species: Archae - 0; Bacteria - 7; Metazoa - 139; Fungi - 93; Plants - 32; Viruses - 0; Other Eukaryotes - 99 (source: NCBI BLink). |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.32901.1.A1_at | 0.0742842 | 1.441793 | Glyma06g19710.2 | AT5G50100.1 | | FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: mitochondrion; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Thioredoxin-like fold (InterPro:IPR012336), Putative thiol-disulphide oxidoreductase DCC (InterPro:IPR007263); Has 598 Blast hits to 598 proteins in 152 species: Archae - 0; Bacteria - 266; Metazoa - 3; Fungi - 0; Plants - 27; Viruses - 0; Other Eukaryotes - 302 (source: NCBI BLink). |
| GmaAffx.38475.1.S1_at | 0.0132164 | 1.43864 | Glyma19g41880.2 | AT1G56260.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; Has 13 Blast hits to 13 proteins in 7 species: Archae - 0; Bacteria - 0; Metazoa - 2; Fungi - 0; Plants - 11; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| GmaAffx.25785.1.A1_at | 0.0684677 | 1.438522 | Glyma09g33970.1 | AT2G35470.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; Has 16 Blast hits to 16 proteins in 6 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 16; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| Gma.5149.1.A1_at | 0.0938293 | 1.423706 | Glyma08g22110.1 | AT1G79975.1 | | similar to Os01g0593600 [Oryza sativa (japonica cultivar-group)] (GB:NP_001043462.1); similar to hypothetical protein DDBDRAFT_0190211 [Dictyostelium discoideum AX4] (GB:XP_645912.1); contains domain SUBFAMILY NOT NAMED (PTHR12952:SF5); contains domain FAMILY NOT NAMED (PTHR12952) |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.93149.1.S1_s_at | 0.0623842 | 1.413277 | Glyma16g18370.1 | AT5G11680.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cytosol, plasma membrane; EXPRESSED IN: 26 plant structures; EXPRESSED DURING: 15 growth stages; Has 136 Blast hits to 136 proteins in 70 species: Archae - 0; Bacteria - 0; Metazoa - 58; Fungi - 26; Plants - 24; Viruses - 0; Other Eukaryotes - 28 (source: NCBI BLink). |
| GmaAffx.1301.108.A1_at | 0.0560276 | 1.397143 | Glyma12g37080.1 | AT5G14240.1 | | FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Thioredoxin fold (InterPro:IPR012335), Phosducin (InterPro:IPR001200), Thioredoxin-like fold (InterPro:IPR012336); Has 750 Blast hits to 750 proteins in 282 species: Archae - 0; Bacteria - 2; Metazoa - 486; Fungi - 126; Plants - 51; Viruses - 0; Other Eukaryotes - 85 (source: NCBI BLink). |
| Gma.17778.1.A1_at | 0.0822382 | 1.39006 | Glyma02g05950.1 | AT4G36440.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: endomembrane system; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; Has 16 Blast hits to 16 proteins in 6 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 16; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| Gma.360.1.S1_at | 0.0994872 | 1.383056 | Glyma17g02450.1 | AT5G18850.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: endomembrane system; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; Has 8 Blast hits to 8 proteins in 4 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 8; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| Gma.8725.1.S1_at | 0.0596157 | 1.359324 | Ambiguous Hit | NO BLASTX HIT (E<10-6) | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.3126.1.S1_a_at | 0.0671638 | 1.348637 | Glyma15g17900.2 | AT1G78890.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT1G16840.4); Has 51 Blast hits to 51 proteins in 9 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 51; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| Gma.13853.1.A1_at | 0.0714804 | 1.333097 | Glyma08g07230.1 | AT4G33625.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: plasma membrane, vacuole; EXPRESSED IN: cultured cell; Has 21 Blast hits to 21 proteins in 9 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 21; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| Gma.17131.1.S1_at | 0.0771246 | 1.330271 | Glyma18g16270.1 | AT1G69840.6 | | band 7 family protein; INVOLVED IN: biological_process unknown; LOCATED IN: plasma membrane; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Band 7 protein (InterPro:IPR001107); BEST Arabidopsis thaliana protein match is: band 7 family protein (TAIR:AT5G62740.1); Has 3729 Blast hits to 3728 proteins in 1086 species: Archae - 110; Bacteria - 2212; Metazoa - 272; Fungi - 168; Plants - 160; Viruses - 5; Other Eukaryotes - 802 (source: NCBI BLink). |
| GmaAffx.63919.1.S1_at | 0.0429772 | 1.327301 | Glyma08g20410.1 | AT1G72690.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT5G57000.2); Has 29 Blast hits to 23 proteins in 4 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 29; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.12492.1.S1_at | 0.0842249 | 1.310634 | Glyma08g15270.2 | AT4G09830.1 | | FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Uncharacterised conserved protein UCP009193 (InterPro:IPR016549); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT5G64780.1); Has 56 Blast hits to 56 proteins in 10 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 56; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| GmaAffx.25674.1.S1_at | 0.0641055 | 1.300444 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| Gma.7573.1.S1_at | 0.0762235 | 1.29161 | Glyma07g18180.1 | AT5G06350.1 | | binding; FUNCTIONS IN: binding; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; CONTAINS InterPro DOMAIN/s: Armadillo-type fold (InterPro:IPR016024); BEST Arabidopsis thaliana protein match is: binding (TAIR:AT5G27010.1); Has 221 Blast hits to 213 proteins in 103 species: Archae - 0; Bacteria - 0; Metazoa - 100; Fungi - 83; Plants - 22; Viruses - 0; Other Eukaryotes - 16 (source: NCBI BLink). |
| GmaAffx.11180.1.A1_at | 0.07755 | 1.265209 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| Gma.7858.1.A1_at | 0.0925148 | 1.264813 | Glyma20g35300.2 | AT1G04230.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 14 growth stages; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT5G43720.1); Has 1140 Blast hits to 725 proteins in 146 species: Archae - 0; Bacteria - 18; Metazoa - 635; Fungi - 256; Plants - 52; Viruses - 5; Other Eukaryotes - 174 (source: NCBI BLink). |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.87331.1.S1_at | 0.0658173 | 1.246807 | Glyma03g34590.1 | AT1G03910.1 | | EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Cactin, central region (InterPro:IPR018816); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT2G36815.2); Has 11516 Blast hits to 6722 proteins in 356 species: Archae - 23; Bacteria - 259; Metazoa - 6122; Fungi - 1009; Plants - 493; Viruses - 33; Other Eukaryotes - 3577 (source: NCBI BLink). |
| GmaAffx.44856.1.A1_at | 0.0475936 | 1.208017 | Glyma11g08650.2 | AT1G05970.2 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; Has 43 Blast hits to 43 proteins in 13 species: Archae - 0; Bacteria - 0; Metazoa - 2; Fungi - 0; Plants - 41; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| GmaAffx.78679.1.S1_at | 0.0339229 | -11.1491 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| GmaAffx.21429.1.A1_at | 0.0500445 | -9.41925 | Glyma01g02510.1 | NO BLASTP HIT (E<10-6) | | |
| Gma.13734.1.A1_at | 0.0267886 | -7.07176 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| GmaAffx.57799.1.S1_at | 0.0954276 | -6.98919 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| GmaAffx.78077.1.S1_at | 0.0085675 | -6.06667 | Glyma03g37430.1 | NO BLASTP HIT (E<10-6) | | |
| GmaAffx.78635.1.S1_at | 0.0795603 | -5.86433 | Glyma07g04000.1 | NO BLASTP HIT (E<10-6) | | |
| GmaAffx.52369.2.S1_at | 0.0689596 | -5.80708 | Glyma19g06110.1 | AT5G38690.1 | | FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: DDT superfamily (InterPro:IPR018501), DDT subgroup (InterPro:IPR018500), Cell division cycle-associated protein (InterPro:IPR018866); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT1G67780.1); Has 492 Blast hits to 468 proteins in 97 species: Archae - 9; Bacteria - 41; Metazoa - 179; Fungi - 19; Plants - 99; Viruses - 2; Other Eukaryotes - 143 (source: NCBI BLink). |
| Gma.16832.1.S1_s_at | 0.0367848 | -5.60333 | Glyma02g09210.1 | NO BLASTP HIT (E<10-6) | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.10830.1.S1_at | 0.0438598 | -5.59317 | Glyma10g28680.1 | AT1G09310.1 | | unknown protein; INVOLVED IN: biological_process unknown; LOCATED IN: apoplast; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF538 (InterPro:IPR007493); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT1G56580.1); Has 197 Blast hits to 195 proteins in 15 species: Archae - 0; Bacteria - 2; Metazoa - 0; Fungi - 4; Plants - 191; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| GmaAffx.53717.1.S1_at | 0.0370705 | -4.84127 | Glyma10g36150.1 | AT1G71970.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: chloroplast; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT1G22680.1); Has 38 Blast hits to 38 proteins in 9 species: Archae - 0; Bacteria - 0; Metazoa - 3; Fungi - 0; Plants - 30; Viruses - 0; Other Eukaryotes - 5 (source: NCBI BLink). |
| GmaAffx.75094.1.A1_at | 0.0507348 | -4.42279 | Glyma15g14070.1 | NO BLASTP HIT (E<10-6) | | |
| GmaAffx.24931.1.S1_at | 0.0952912 | -4.07832 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| GmaAffx.31311.2.S1_at | 0.0872861 | -3.96733 | Glyma12g31930.1 | NO BLASTP HIT (E<10-6) | | |
| GmaAffx.80463.1.S1_at | 0.0416405 | -3.9653 | Glyma08g00950.1 | NO BLASTP HIT (E<10-6) | | |
| Gma.4843.1.A1_at | 0.0247239 | -3.87973 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| GmaAffx.43236.1.A1_at | 0.072104 | -3.84127 | Glyma10g28850.1 | AT2G24100.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT4G30780.1); Has 44 Blast hits to 44 proteins in 11 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 39; Viruses - 0; Other Eukaryotes - 5 (source: NCBI BLink). |
| Gma.5947.1.S1_s_at | 0.07755 | -3.76954 | Glyma09g04530.1 | NO BLASTP HIT (E<10-6) | | |
| Gma.4964.2.A1_at | 0.0771246 | -3.76363 | Glyma03g37070.1 | NO BLASTP HIT (E<10-6) | | |
| GmaAffx.86022.1.S1_at | 0.0623842 | -3.63774 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| GmaAffx.72807.1.S1_at | 0.097011 | -3.61239 | Glyma08g28050.1 | NO BLASTP HIT (E<10-6) | | |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.77420.1.S1_at | 0.0623842 | -3.4406 | Glyma11g27050.1 | AT1G07220.1 | FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: mitochondrion; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Lipopolysaccharide-modifying protein (InterPro:IPR006598), Protein of unknown function DUF821, CAP10-like (InterPro:IPR008539); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT5G23850.1); Has 456 Blast hits to 452 proteins in 94 species: Archae - 0; Bacteria - 34; Metazoa - 198; Fungi - 90; Plants - 113; Viruses - 2; Other Eukaryotes - 19 (source: NCBI BLink). |
| GmaAffx.18019.1.S1_at | 0.0684677 | -3.4306 | Glyma17g24930.1 | AT5G10580.1 | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: endomembrane system; EXPRESSED IN: hypocotyl, flower, root; EXPRESSED DURING: petal differentiation and expansion stage; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF599 (InterPro:IPR006747); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT4G31330.1); Has 253 Blast hits to 252 proteins in 83 species: Archae - 0; Bacteria - 138; Metazoa - 0; Fungi - 0; Plants - 95; Viruses - 0; Other Eukaryotes - 20 (source: NCBI BLink). |
| GmaAffx.51714.1.A1_at | 0.0267886 | -3.37487 | Glyma15g12310.1 | NO BLASTP HIT (E<10-6) | |
| GmaAffx.12086.1.S1_at | 0.0964098 | -3.36052 | Glyma10g05330.1 | AT3G09570.1 | FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: endomembrane system, integral to membrane; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Transmembrane receptor, eukaryota (InterPro:IPR009637); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT5G18520.1); Has 385 Blast hits to 385 proteins in 83 species: Archae - 0; Bacteria - 2; Metazoa - 216; Fungi - 12; Plants - 132; Viruses - 0; Other Eukaryotes - 23 (source: NCBI BLink). |
| GmaAffx.61499.1.A1_at | 0.0969491 | -3.28975 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.83107.1.S1_at | 0.085144 | -3.24418 | Glyma16g15790.1 | AT2G26570.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: plasma membrane; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF827, plant (InterPro:IPR008545); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT4G33390.1); Has 116898 Blast hits to 58103 proteins in 2191 species: Archae - 1417; Bacteria - 21958; Metazoa - 54071; Fungi - 9153; Plants - 5267; Viruses - 512; Other Eukaryotes - 24520 (source: NCBI BLink). |
| Gma.5141.1.S1_at | 0.0689596 | -3.17552 | Glyma14g37810.1 | AT1G21280.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; Has 594 Blast hits to 592 proteins in 21 species: Archae - 0; Bacteria - 0; Metazoa - 4; Fungi - 0; Plants - 590; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| GmaAffx.85693.1.S1_at | 0.0996488 | -3.16813 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| Gma.13092.1.S1_a_at | 0.0595153 | -3.12027 | Glyma14g08270.3 | AT1G47740.2 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF862, eukaryotic (InterPro:IPR008580); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT5G25170.1); Has 559 Blast hits to 559 proteins in 131 species: Archae - 0; Bacteria - 2; Metazoa - 186; Fungi - 58; Plants - 182; Viruses - 0; Other Eukaryotes - 131 (source: NCBI BLink). |
| GmaAffx.17308.1.A1_at | 0.0500445 | -3.05137 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.12905.1.A1_at | 0.0537759 | -3.01086 | Glyma03g27760.2 | AT5G07050.1 | | LOCATED IN: membrane; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF6, transmembrane (InterPro:IPR000620); BEST Arabidopsis thaliana protein match is: nodulin MtN21 family protein (TAIR:AT2G40900.1); Has 2288 Blast hits to 2270 proteins in 280 species: Archae - 23; Bacteria - 774; Metazoa - 5; Fungi - 0; Plants - 641; Viruses - 0; Other Eukaryotes - 845 (source: NCBI BLink). |
| GmaAffx.23647.1.S1_at | 0.0557173 | -2.93556 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| GmaAffx.16959.1.S1_s_at | 0.0714804 | -2.91725 | Glyma13g35050.1 | AT5G53880.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: mitochondrion, plasma membrane; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; Has 1379 Blast hits to 641 proteins in 104 species: Archae - 20; Bacteria - 31; Metazoa - 673; Fungi - 150; Plants - 69; Viruses - 10; Other Eukaryotes - 426 (source: NCBI BLink). |
| GmaAffx.37994.1.A1_at | 0.0872493 | -2.8627 | Glyma04g03510.1 | AT5G65650.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF1195 (InterPro:IPR010608); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT4G36660.1); Has 43 Blast hits to 43 proteins in 11 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 43; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| GmaAffx.61492.1.A1_at | 0.0267886 | -2.82926 | Glyma01g38770.1 | NO BLASTP HIT (E<10-6) | | |
| GmaAffx.55786.1.S1_at | 0.0392765 | -2.7897 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| GmaAffx.56178.1.A1_at | 0.0394748 | -2.71917 | Glyma13g35660.1 | NO BLASTP HIT (E<10-6) | | |
| GmaAffx.45984.1.A1_at | 0.0595153 | -2.71056 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |

| | | | | | |
|---|---|---|---|---|---|
| Gma.16660.1.S1_at | 0.0581114 | -2.6848 | Glyma18g02660.1 | AT3G56140.1 | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: chloroplast thylakoid lumen, chloroplast; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF399 (InterPro:IPR007314); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT2G40400.2); Has 354 Blast hits to 354 proteins in 86 species: Archae - 0; Bacteria - 111; Metazoa - 36; Fungi - 6; Plants - 172; Viruses - 6; Other Eukaryotes - 23 (source: NCBI BLink). |
| Gma.13543.1.A1_at | 0.0649873 | -2.64895 | Glyma13g20850.1 | AT5G27730.1 | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 14 growth stages; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:ATSG47900.1); Has 651 Blast hits to 616 proteins in 129 species: Archae - 0; Bacteria - 235; Metazoa - 100; Fungi - 0; Plants - 60; Viruses - 0; Other Eukaryotes - 256 (source: NCBI BLink). |
| GmaAffx.80463.2.S1_at | 0.066588 | -2.63759 | Glyma08g00950.1 | NO BLASTP HIT (E<10-6) | |
| GmaAffx.40337.1.S1_at | 0.0806207 | -2.61914 | Glyma05g22090.1 | NO BLASTP HIT (E<10-6) | |
| GmaAffx.84038.1.A1_at | 0.066588 | -2.55791 | Glyma18g46040.1 | AT5G01150.1 | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF674 (InterPro:IPR007750); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT5G01130.1); Has 188 Blast hits to 125 proteins in 8 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 188; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| GmaAffx.6624.1.A1_at | 0.0581114 | -2.54217 | Glyma05g08350.1 | NO BLASTP HIT (E<10-6) | |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.15786.1.S1_at | 0.097933 | -2.47329 | Glyma13g17140.1 | | AT1G19130.1 | FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF985 (InterPro:IPR009327), RmlC-like jelly roll fold (InterPro:IPR014710); Has 778 Blast hits to 778 proteins in 280 species: Archae - 8; Bacteria - 513; Metazoa - 2; Fungi - 16; Plants - 22; Viruses - 0; Other Eukaryotes - 217 (source: NCBI BLink). |
| GmaAffx.7416.1.A1_at | 0.0731769 | -2.42942 | Glyma04g02350.1 | | AT1G76020.1 | INVOLVED IN: biological_process unknown; LOCATED IN: cell wall, vacuole; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Thioredoxin fold (InterPro:IPR012335), Thioredoxin-like fold (InterPro:IPR012336); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT1G20225.1); Has 118 Blast hits to 116 proteins in 37 species: Archae - 0; Bacteria - 51; Metazoa - 12; Fungi - 4; Plants - 28; Viruses - 0; Other Eukaryotes - 23 (source: NCBI BLink). |
| Gma.4901.1.S1_at | 0.0921169 | -2.40612 | Glyma17g33500.2 | | AT1G72510.2 | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF1677, plant (InterPro:IPR012876); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT2G09970.1); Has 130 Blast hits to 129 proteins in 7 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 130; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| Gma.6416.1.S1_at | 0.0925065 | -2.38339 | Glyma08g27380.1 | | NO BLASTP HIT (E<10-6) | |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.2281.1.S1_at | 0.0471957 | -2.38143 | Glyma03g40260.2 | AT3G05010.1 | | transmembrane protein, putative; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: membrane; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function, transmembrane-40 (InterPro:IPR018781); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT5G27210.1); Has 103 Blast hits to 103 proteins in 34 species: Archae - 0; Bacteria - 4; Metazoa - 71; Fungi - 0; Plants - 27; Viruses - 0; Other Eukaryotes - 1 (source: NCBI BLink). |
| GmaAffx.1956.1.S1_at | 0.0957992 | -2.36872 | Glyma02g47770.1 | AT2G22795.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT4G37820.1); Has 742033 Blast hits to 185230 proteins in 3414 species: Archae - 3102; Bacteria - 91104; Metazoa - 304818; Fungi - 76748; Plants - 32809; Viruses 4044; Other Eukaryotes - 229408 (source: NCBI BLink). |
| GmaAffx.15149.1.A1_at | 0.0812658 | -2.3541 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| GmaAffx.40015.1.S1_at | 0.0336887 | -2.32071 | Glyma07g39370.1 | AT4G22190.1 | | unknown protein; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; Has 261 Blast hits to 128 proteins in 36 species: Archae - 0; Bacteria - 2; Metazoa - 26; Fungi - 10; Plants - 24; Viruses - 0; Other Eukaryotes - 199 (source: NCBI BLink). |
| GmaAffx.57983.1.S1_at | 0.0981988 | -2.31616 | Glyma06g18950.1 | NO BLASTP HIT (E<10-6) | | |
| Gma.14570.2.S1_at | 0.0580906 | -2.29324 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| GmaAffx.15664.1.S1_at | 0.0430652 | -2.28277 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.6022.1.S1_at | 0.0581114 | -2.27808 | Glyma15g01020.1 | AT3G15810.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF567 (InterPro:IPR007612); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT1G80120.1); Has 208 Blast hits to 208 proteins in 14 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 7; Plants - 201; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| Gma.8385.1.A1_at | 0.070119 | -2.2745 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| GmaAffx.1620.3.S1_s_at | 0.0416405 | -2.14053 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| GmaAffx.73185.1.S1_at | 0.0815612 | -2.11397 | Glyma13g06460.1 | NO BLASTP HIT (E<10-6) | | |
| Gma.7013.4.S1_a_at | 0.0558816 | -2.10835 | Glyma19g25070.1 | AT5G20090.2 | | unknown protein; INVOLVED IN: biological_process unknown; LOCATED IN: plasma membrane, membrane; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 16 growth stages; CONTAINS InterPro DOMAIN/s: Uncharacterised protein family UPF0041 (InterPro:IPR005336); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT4G22310.1); Has 657 Blast hits to 656 proteins in 149 species: Archae - 0; Bacteria - 0; Metazoa - 314; Fungi - 175; Plants - 98; Viruses - 0; Other Eukaryotes - 70 (source: NCBI BLink). |
| GmaAffx.64943.2.S1_at | 0.0598854 | -2.05967 | Glyma02g06820.2 | AT4G37235.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: endomembrane system; CONTAINS InterPro DOMAIN/s: Uncharacterised protein family UPF0497, trans-membrane plant (InterPro:IPR006702); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT3G50810.1); Has 97 Blast hits to 97 proteins in 8 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 97; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.11221.1.S1_at | 0.0665125 | -2.0593 | No Soybean Match Iden | AT5G12470.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: mitochondrion, chloroplast, plastid, chloroplast inner membrane, chloroplast envelope; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 14 growth stages; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT2G40400.2); Has 570 Blast hits to 490 proteins in 89 species: Archae - 0; Bacteria - 92; Metazoa - 134; Fungi - 23; Plants - 222; Viruses - 27; Other Eukaryotes - 72 (source: NCBI BLink). |
| Gma.10240.1.A1_at | 0.0267886 | -2.04809 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| Gma.12888.1.A1_s_at | 0.07755 | -2.02896 | Glyma17g35230.1 | AT4G33800.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 12 plant structures; EXPRESSED DURING: 7 growth stages; Has 27 Blast hits to 27 proteins in 5 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 27; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| Gma.14190.1.A1_at | 0.0517154 | -2.01312 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| Gma.477.1.S1_at | 0.0629145 | -1.99469 | Glyma15g02810.1 | AT1G17620.1 | | FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: plasma membrane; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Harpin-induced 1 (InterPro:IPR010847); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT5G11890.1); Has 539 Blast hits to 538 proteins in 22 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 539; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| GmaAffx.1301.32.A1_at | 0.0808295 | -1.98945 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| GmaAffx.83877.1.S1_at | 0.0595153 | -1.98385 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.6132.1.S1_at | 0.0894786 | -1.95075 | Ambiguous Hit | | AT5G52980.1 | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 14 growth stages; Has 18 Blast hits to 18 proteins in 10 species: Archae - 0; Bacteria - 0; Metazoa - 2; Fungi - 0; Plants - 13; Viruses - 0; Other Eukaryotes - 3 (source: NCBI BLink). |
| Gma.18072.1.S1_at | 0.0930231 | -1.93073 | Ambiguous Hit | | AT3G14920.1 | unknown protein; INVOLVED IN: biological_process unknown; LOCATED IN: plasma membrane, vacuole, plant-type cell wall; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT5G05480.1); Has 144 Blast hits to 133 proteins in 54 species: Archae - 12; Bacteria - 7; Metazoa - 0; Fungi - 71; Plants - 51; Viruses - 0; Other Eukaryotes - 3 (source: NCBI BLink). |
| GmaAffx.31210.1.S1_at | 0.0577287 | -1.91689 | Glyma18g38620.1 | | AT5G02440.1 | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 17 plant structures; EXPRESSED DURING: 11 growth stages; Has 29 Blast hits to 29 proteins in 8 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 29; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| GmaAffx.15318.1.S1_at | 0.0560276 | -1.90946 | Glyma19g33370.3 | | AT1G20770.1 | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 14 growth stages; Has 37 Blast hits to 37 proteins in 14 species: Archae - 0; Bacteria - 0; Metazoa - 19; Fungi - 2; Plants - 16; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| Gma.17444.1.S1_at | 0.0581114 | -1.89996 | No Soybean Match Iden | | NO BLASTX HIT (E<10-6) | |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.13382.1.A1_at | 0.0636582 | -1.8923 | Glyma20g37770.1 | AT5G27560.1 | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: chloroplast; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; Has 227 Blast hits to 227 proteins in 62 species: Archae - 0; Bacteria - 92; Metazoa - 0; Fungi - 0; Plants - 34; Viruses - 0; Other Eukaryotes - 101 (source: NCBI BLink). |
| GmaAffx.56363.1.S1_at | 0.0962248 | -1.89128 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| Gma.5677.1.S1_at | 0.0337667 | -1.88324 | Glyma09g14860.1 | AT4G30996.1 | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: endomembrane system; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF1068 (InterPro:IPR010471); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT2G24290.1); Has 51 Blast hits to 51 proteins in 9 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 50; Viruses - 0; Other Eukaryotes - 1 (source: NCBI BLink). |
| Gma.4752.1.S1_s_at | 0.0595153 | -1.88271 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| GmaAffx.7258.1.S1_s_at | 0.066588 | -1.87784 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| GmaAffx.79526.1.A1_at | 0.0649873 | -1.87107 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| Gma.2281.1.S1_x_at | 0.0595153 | -1.8701 | Glyma03g40260.2 | AT3G05010.1 | transmembrane protein, putative; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: membrane; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function, transmembrane-40 (InterPro:IPR018781); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT5G27210.1); Has 103 Blast hits to 103 proteins in 34 species: Archae - 0; Bacteria - 4; Metazoa - 71; Fungi - 0; Plants - 27; Viruses - 0; Other Eukaryotes - 1 (source: NCBI BLink). |
| Gma.11097.1.A1_at | 0.0694153 | -1.86763 | Glyma17g12030.1 | NO BLASTP HIT (E<10-6) | | |
| Gma.12915.1.A1_s_at | 0.060918 | -1.84158 | Glyma09g37030.1 | NO BLASTP HIT (E<10-6) | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.7332.1.S1_at | 0.0231573 | -1.83894 | Glyma03g35780.1 | AT5G06270.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: chloroplast; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 11 growth stages; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT3G11600.1); Has 108 Blast hits to 108 proteins in 10 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 108; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| GmaAffx.54698.1.S1_at | 0.0581114 | -1.81724 | No Soybean Match Iden | AT4G27810.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT5G53030.1); Has 40 Blast hits to 35 proteins in 5 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 40; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| Gma.13704.1.A1_s_at | 0.0955455 | -1.81067 | Glyma02g44520.1 | NO BLASTP HIT (E<10-6) | | |
| GmaAffx.74716.1.A1_at | 0.0753187 | -1.80484 | Glyma12g06420.1 | NO BLASTP HIT (E<10-6) | | |
| Gma.2823.1.S1_at | 0.0688506 | -1.80409 | Glyma08g10990.1 | AT1G27385.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: chloroplast; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF493 (InterPro:IPR007454); Has 46 Blast hits to 46 proteins in 13 species: Archae - 0; Bacteria - 4; Metazoa - 0; Fungi - 0; Plants - 41; Viruses - 0; Other Eukaryotes - 1 (source: NCBI BLink). |
| Gma.13633.1.A1_at | 0.0702895 | -1.78368 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.4509.1.S1_at | 0.0842249 | -1.74203 | Glyma11g06440.1 | AT4G37090.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; Has 1174 Blast hits to 794 proteins in 120 species: Archae - 2; Bacteria - 24; Metazoa - 344; Fungi - 98; Plants - 43; Viruses - 11; Other Eukaryotes - 652 (source: NCBI BLink). |
| GmaAffx.46382.1.S1_at | 0.0558816 | -1.70399 | Glyma20g29340.1 | NO BLASTP HIT (E<10-6) | | |
| GmaAffx.73401.1.S1_x_at | 0.0641208 | -1.70128 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| GmaAffx.31036.1.S1_at | 0.0839752 | -1.70098 | Glyma11g06120.1 | AT5G66820.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; Has 31 Blast hits to 31 proteins in 15 species: Archae - 0; Bacteria - 2; Metazoa - 1; Fungi - 4; Plants - 7; Viruses - 2; Other Eukaryotes - 15 (source: NCBI BLink). |
| Gma.16218.1.A1_at | 0.0967103 | -1.69948 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| GmaAffx.29964.1.A1_at | 0.0968339 | -1.68198 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| GmaAffx.36996.1.S1_at | 0.0771246 | -1.67748 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| Gma.9228.1.S1_at | 0.0522596 | -1.67483 | Glyma13g24380.1 | AT1G28400.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: endomembrane system; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT2G33850.1); Has 44658 Blast hits to 17636 proteins in 467 species: Archae - 50; Bacteria - 837; Metazoa - 1245; Fungi - 895; Plants - 135; Viruses - 63; Other Eukaryotes - 41433 (source: NCBI BLink). |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.17283.1.A1_at | 0.0641208 | -1.66655 | Glyma01g34820.1 | AT4G28910.3 | | LOCATED IN: cellular_component unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF1675 (InterPro:IPR012463); BEST Arabidopsis thaliana protein match is: nuclear transport factor 2 (NTF2) family protein / RNA recognition motif (RRM)-containing protein (TAIR:AT3G07250.1); Has 223 Blast hits to 204 proteins in 56 species: Archae - 2; Bacteria - 11; Metazoa - 48; Fungi - 32; Plants - 89; Viruses - 3; Other Eukaryotes - 38 (source: NCBI BLink). |
| Gma.12921.1.A1_at | 0.07755 | -1.65117 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| GmaAffx.29955.1.S1_at | 0.0594787 | -1.63994 | Glyma15g07410.1 | NO BLASTP HIT (E<10-6) | | |
| Gma.16421.1.S1_at | 0.0935144 | -1.63035 | Glyma08g04240.1 | NO BLASTP HIT (E<10-6) | | |
| GmaAffx.7738.1.S1_s_at | 0.0736827 | -1.62968 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| GmaAffx.91720.1.S1_s_at | 0.0637147 | -1.62185 | Glyma05g31830.1 | AT2G21870.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: mitochondrion, nucleolus, chloroplast, plastid, membrane; EXPRESSED IN: 27 plant structures; EXPRESSED DURING: 16 growth stages; Has 84 Blast hits to 83 proteins in 36 species: Archae - 2; Bacteria - 11; Metazoa - 21; Fungi - 0; Plants - 30; Viruses - 0; Other Eukaryotes - 20 (source: NCBI BLink). |
| Gma.907.1.A1_at | 0.0475936 | -1.60126 | Glyma02g06360.1 | AT5G49560.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 16 plant structures; EXPRESSED DURING: 8 growth stages; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT3G50850.1); Has 869 Blast hits to 869 proteins in 147 species: Archae - 0; Bacteria - 25; Metazoa - 384; Fungi - 266; Plants - 132; Viruses - 0; Other Eukaryotes - 62 (source: NCBI BLink). |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.17465.1.S1_at | 0.0560276 | -1.59839 | Glyma02g35180.1 | AT3G44150.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: plasma membrane; EXPRESSED IN: cultured cell; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT3G11800.1); Has 46 Blast hits to 44 proteins in 11 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 45; Viruses - 0; Other Eukaryotes - 1 (source: NCBI BLink). |
| Gma.3660.1.S1_at | 0.0998527 | -1.59164 | Glyma20g24750.4 | AT1G68650.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: membrane; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Uncharacterised protein family UPF0016 (InterPro:IPR001727); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT1G25520.1); Has 1272 Blast hits to 1209 proteins in 450 species: Archae - 18; Bacteria - 687; Metazoa - 131; Fungi - 100; Plants - 104; Viruses - 0; Other Eukaryotes - 232 (source: NCBI BLink). |
| Gma.4539.1.A1_s_at | 0.0742842 | -1.59075 | Glyma10g08640.1 | AT3G52610.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 15 growth stages; Has 38 Blast hits to 37 proteins in 11 species: Archae - 0; Bacteria - 5; Metazoa - 0; Fungi - 0; Plants - 31; Viruses - 0; Other Eukaryotes - 2 (source: NCBI BLink). |
| GmaAffx.85497.1.S1_at | 0.0600598 | -1.58974 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| Gma.13692.1.A1_s_at | 0.0517137 | -1.58513 | Ambiguous Hit | NO BLASTX HIT (E<10-6) | | |
| Gma.4164.3.S1_s_at | 0.0972565 | -1.54743 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| Gma.17244.1.A1_at | 0.0968935 | -1.5315 | Glyma16g02380.1 | NO BLASTP HIT (E<10-6) | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.5764.1.S1_at | 0.0601266 | -1.51955 | Glyma09g35650.1 | AT3G01130.1 | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: endomembrane system; EXPRESSED IN: male gametophyte, pollen tube; EXPRESSED DURING: M germinated pollen stage; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT5G15320.1); Has 37 Blast hits to 37 proteins in 7 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 37; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| GmaAffx.20876.1.S1_at | 0.0671737 | -1.51452 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| GmaAffx.50652.1.A1_s_at | 0.066588 | -1.50352 | Glyma08g04040.1 | AT4G28840.1 | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: chloroplast; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT2G20080.1); Has 40 Blast hits to 38 proteins in 9 species: Archae - 0; Bacteria - 0; Metazoa - 2; Fungi - 0; Plants - 36; Viruses - 0; Other Eukaryotes - 2 (source: NCBI BLink). |
| GmaAffx.73658.1.S1_s_at | 0.0623842 | -1.49775 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| Gma.13940.1.S1_at | 0.0893205 | -1.4969 | Glyma15g14530.1 | AT3G17300.1 | EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Complex 1 LYR protein (InterPro:IPR008011); Has 28 Blast hits to 28 proteins in 12 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 23; Viruses - 0; Other Eukaryotes - 5 (source: NCBI BLink). |
| Gma.1356.2.S1_a_at | 0.089272 | -1.48853 | Glyma18g46830.1 | AT2G45990.2 | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: chloroplast; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; Has 251 Blast hits to 251 proteins in 66 species: Archae - 0; Bacteria - 105; Metazoa - 0; Fungi - 0; Plants - 34; Viruses - 0; Other Eukaryotes - 112 (source: NCBI BLink). |
| Gma.621.1.A1_at | 0.0577786 | -1.48592 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |

| | | | | | |
|---|---|---|---|---|---|
| Gma.16418.1.A1_at | 0.0893658 | -1.47619 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | |
| Gma.8759.1.A1_at | 0.0446585 | -1.46976 | Glyma13g17360.2 | AT1G19240.1 | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; Has 11 Blast hits to 11 proteins in 5 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 11; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| GmaAffx.3715.1.S1_at | 0.0998527 | -1.4636 | Glyma13g41730.2 | AT5G12900.1 | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: chloroplast; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 13 growth stages; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT1G12330.1); Has 278 Blast hits to 272 proteins in 46 species: Archae - 0; Bacteria - 26; Metazoa - 55; Fungi - 10; Plants - 146; Viruses - 0; Other Eukaryotes - 41 (source: NCBI BLink). |
| GmaAffx.66263.1.S1_at | 0.0625053 | -1.44275 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | |
| Gma.3482.1.S1_at | 0.0565355 | -1.42164 | Glyma01g44500.1 | AT4G14500.1 | FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Lipid-binding START (InterPro:IPR002913); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT3G23080.1); Has 257 Blast hits to 256 proteins in 63 species: Archae - 0; Bacteria - 0; Metazoa - 167; Fungi - 0; Plants - 83; Viruses - 0; Other Eukaryotes - 7 (source: NCBI BLink). |
| Gma.16600.2.A1_at | 0.0848845 | -1.40845 | Glyma10g06070.1 | AT2G37195.1 | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: endomembrane system; Has 12 Blast hits to 12 proteins in 6 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 12; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.15923.1.S1_at | 0.0416405 | -1.4063 | Glyma04g35050.1 | AT5G62575.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: mitochondrion; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 15 growth stages; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT3G47833.1); Has 26 Blast hits to 26 proteins in 8 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 26; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| GmaAffx.86989.1.A1_at | 0.0979683 | -1.39851 | Glyma19g33070.1 | AT2G42040.1 | | FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: WRC (InterPro:IPR014977); Has 154 Blast hits to 154 proteins in 12 species: Archae - 0; Bacteria - 0; Metazoa - 4; Fungi - 0; Plants - 149; Viruses - 0; Other Eukaryotes - 1 (source: NCBI BLink). |
| Gma.9324.1.S1_at | 0.0813937 | -1.388 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| Gma.2590.2.S1_at | 0.0510522 | -1.38298 | Glyma01g38210.9 | NO BLASTP HIT (E<10-6) | | |
| GmaAffx.44838.1.A1_s_at | 0.0500445 | -1.32795 | Glyma03g38310.1 | AT3G62140.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 12 growth stages; Has 331 Blast hits to 328 proteins in 104 species: Archae - 0; Bacteria - 6; Metazoa - 181; Fungi - 52; Plants - 23; Viruses - 0; Other Eukaryotes - 69 (source: NCBI BLink). |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.12479.1.S1_at | 0.0921169 | -1.32364 | Glyma01g45250.1 | AT2G14660.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 17 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Uncharacterised protein family UPF0310 (InterPro:IPR002740); Has 2334 Blast hits to 2333 proteins in 479 species: Archae - 5; Bacteria - 781; Metazoa - 60; Fungi - 50; Plants - 22; Viruses - 0; Other Eukaryotes - 1416 (source: NCBI BLink). |
| GmaAffx.10915.1.S1_at | 0.0776647 | -1.31294 | Ambiguous Hit | NO BLASTX HIT (E<10-6) | | |
| GmaAffx.15170.1.S1_at | 0.0913044 | -1.27682 | Glyma14g34800.1 | NO BLASTP HIT (E<10-6) | | |
| GmaAffx.50791.1.S1_at | 0.068696 | -1.27416 | Glyma02g07570.1 | AT2G31600.1 | | unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: chloroplast; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT3G53860.1); Has 29 Blast hits to 28 proteins in 8 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 26; Viruses - 0; Other Eukaryotes - 3 (source: NCBI BLink). |
| GmaAffx.13656.1.S1_at | 0.0814926 | -1.27046 | No Soybean Match Iden | NO BLASTX HIT (E<10-6) | | |
| GmaAffx.87925.1.S1_s_at | 0.081224 | -1.23579 | Glyma17g11180.1 | AT1G15860.1 | | INVOLVED IN: biological_process unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Defective in cullin neddylation (InterPro:IPR014764), Protein of unknown function DUF298 (InterPro:IPR005176); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT3G12760.1); Has 576 Blast hits to 576 proteins in 121 species: Archae - 0; Bacteria - 0; Metazoa - 403; Fungi - 78; Plants - 51; Viruses - 0; Other Eukaryotes - 44 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.3520.1.S1_at | 0.0862961 | -1.11648 | Glyma09g32660.3 | AT4G28910.3 | LOCATED IN: cellular_component unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF1675 (InterPro:IPR012463); BEST Arabidopsis thaliana protein match is: nuclear transport factor 2 (NTF2) family protein / RNA recognition motif (RRM)-containing protein (TAIR:AT3G07250.1); Has 223 Blast hits to 204 proteins in 56 species: Archae - 2; Bacteria - 11; Metazoa - 48; Fungi - 32; Plants - 89; Viruses - 3; Other Eukaryotes - 38 (source: NCBI BLink). |

**Table 11 Candidate genes related to transporter**

| Affy ID | Q-value | Fold change | Soybean Gene | Arabidopsis homolog | Description |
|---|---|---|---|---|---|
| Gma.7737.1.S1_at | 0.02697 | 14.7293 | No Soybean Match Identified, E<10E-30, Perecent ID>95% | AT1G15170.1 | MATE efflux family protein; FUNCTIONS IN: drug transporter activity, antiporter activity, transporter activity; INVOLVED IN: multidrug transport; LOCATED IN: chloroplast, membrane; EXPRESSED IN: 17 plant structures; EXPRESSED DURING: 9 growth stages; CONTAINS InterPro DOMAIN/s: MATE family transporter related protein (InterPro:IPR015521), Multi antimicrobial extrusion protein MatE (InterPro:IPR002528); BEST Arabidopsis thaliana protein match is: MATE efflux family protein (TAIR:AT1G15180.1); Has 4882 Blast hits to 4835 proteins in 1070 species: Archae - 55; Bacteria - 2976; Metazoa - 126; Fungi - 208; Plants - 690; Viruses - 0; Other Eukaryotes - 827 (source: NCBI BLink). |
| Gma.6736.1.S1_at | 0.03172 | 4.08722 | Glyma18g02580.1 | AT3G56200.1 | Encodes a putative amino acid transporter. |
| GmaAffx.80684.1.S1_at | 0.02679 | 4.08466 | Glyma04g10550.1 | AT1G28230.1 | Encodes a transporter that transports purines,cytokinins and other adenine derivatives. Expressed in the leaf hydathodes where it may be involved in re-uptake of cytokinins during guttation. |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.19826.1.A1_at | 0.03377 | 4.03586 | Glyma10g42190.1 | AT4G39670.1 | glycolipid binding / glycolipid transporter; FUNCTIONS IN: glycolipid transporter activity, glycolipid binding; INVOLVED IN: glycolipid transport; LOCATED IN: cytoplasm; EXPRESSED IN: 16 plant structures; EXPRESSED DURING: 10 growth stages; CONTAINS InterPro DOMAIN/s: Glycolipid transfer protein, GLTP (InterPro:IPR014830); BEST Arabidopsis thaliana protein match is: ACD11 (ACCELERATED CELL DEATH 11); sphingosine transmembrane transporter (TAIR:AT2G34690.1); Has 190 Blast hits to 190 proteins in 59 species: Archae - 0; Bacteria - 0; Metazoa - 102; Fungi - 9; Plants - 66; Viruses - 0; Other Eukaryotes - 13 (source: NCBI BLink). |
| Gma.12301.1.S1_at | 0.05105 | 3.09713 | Glyma14g36870.1 | AT4G27590.1 | copper-binding protein-related; FUNCTIONS IN: copper ion binding, metal ion binding; INVOLVED IN: copper ion transport, metal ion transport; LOCATED IN: cellular_component unknown; EXPRESSED IN: 8 plant structures; EXPRESSED DURING: 6 growth stages; CONTAINS InterPro DOMAIN/s: Heavy metal transport/detoxification protein (InterPro:IPR006121); BEST Arabidopsis thaliana protein match is: metal ion binding (TAIR:AT3G56891.1); Has 394 Blast hits to 394 proteins in 18 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 394; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). |
| GmaAffx.64895.1.S1_at | 0.06013 | 3.04542 | Glyma07g39760.1 | AT4G16380.2 | LOCATED IN: cellular_component unknown; BEST Arabidopsis thaliana protein match is: metal ion binding (TAIR:AT1G49420.1); Has 3600 Blast hits to 3158 proteins in 308 species: Archae - 4; Bacteria - 309; Metazoa - 1646; Fungi - 490; Plants - 698; Viruses - 60; Other Eukaryotes - 393 (source: NCBI BLink). |

| | | | No Soybean Match Identified, E<10E-30, Perecent ID>95% | ATMG01080.1 | subunit 9 of mitochondrial F0-ATPase |
|---|---|---|---|---|---|
| GmaAffx.11243.1.S1_s_at | 0.06582 | 2.93279 | | | glycerol-3-phosphate transporter, putative / glycerol 3-phosphate permease, putative; FUNCTIONS IN: sugar:hydrogen symporter activity; INVOLVED IN: carbohydrate transport; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Major facilitator superfamily MFS-1 (InterPro:IPR011701), Major facilitator superfamily, general substrate transporter (InterPro:IPR016196); BEST Arabidopsis thaliana protein match is: glycerol-3-phosphate transporter, putative / glycerol 3-phosphate permease, putative (TAIR:AT3G47420.1); Has 12133 Blast hits to 12040 proteins in 1151 species: Archae - 186; Bacteria - 9351; Metazoa - 1144; Fungi - 578; Plants - 112; Viruses - 0; Other Eukaryotes - 762 (source: NCBI BLink). |
| GmaAffx.80671.1.S1_at | 0.04288 | 2.70518 | Glyma09g13170.1 | AT2G13100.1 | permease, putative; FUNCTIONS IN: transmembrane transporter activity; INVOLVED IN: transport; LOCATED IN: cell wall, membrane; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Xanthine/uracil/vitamin C permease (InterPro:IPR006043); BEST Arabidopsis thaliana protein match is: xanthine/uracil permease family protein (TAIR:AT5G49990.1); Has 4980 Blast hits to 4962 proteins in 950 species: Archae - 37; Bacteria - 3560; Metazoa - 307; Fungi - 91; Plants - 296; Viruses - 1; Other Eukaryotes - 688 (source: NCBI BLink). |
| GmaAffx.7387.1.S1_at | 0.02679 | 2.36443 | Glyma18g18060.1 | AT5G62890.3 | |

| | | | | | |
|---|---|---|---|---|---|
| Gma.4431.1.S1_at | 0.05603 | 1.98211 | Glyma04g42250.1 | AT1G21380.1 | VHS domain-containing protein / GAT domain-containing protein; FUNCTIONS IN: protein transporter activity; INVOLVED IN: intracellular protein transport, intra-Golgi vesicle-mediated transport; LOCATED IN: Golgi stack, intracellular; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: VHS (InterPro:IPR002014), Target of Myb protein 1 (InterPro:IPR014645), GAT (InterPro:IPR004152), VHS subgroup (InterPro:IPR018205), ENTH/VHS (InterPro:IPR008942); BEST Arabidopsis thaliana protein match is: VHS domain-containing protein / GAT domain-containing protein (TAIR:AT1G76970.1); Has 1297 Blast hits to 1295 proteins in 140 species: Archae - 0; Bacteria - 6; Metazoa - 767; Fungi - 311; Plants - 152; Viruses - 0; Other Eukaryotes - 61 (source: NCBI BLink). |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.26359.1.S1_at | 0.07755 | 1.93774 | Glyma12g06380.3 | AT5G59250.1 | sugar transporter family protein; FUNCTIONS IN: carbohydrate transmembrane transporter activity, sugar:hydrogen symporter activity; INVOLVED IN: transport, transmembrane transport; LOCATED IN: chloroplast, membrane, chloroplast envelope; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Sugar transporter, conserved site (InterPro:IPR005829), General substrate transporter (InterPro:IPR005828), Sugar/inositol transporter (InterPro:IPR003663), Major facilitator superfamily, general substrate transporter (InterPro:IPR016196); BEST Arabidopsis thaliana protein match is: sugar transporter family protein (TAIR:AT5G17010.3); Has 27429 Blast hits to 27047 proteins in 1483 species: Archae - 420; Bacteria - 15136; Metazoa - 3762; Fungi - 4998; Plants - 1535; Viruses - 2; Other Eukaryotes - 1576 (source: NCBI BLink). |
| Gma.541.1.S1_at | 0.0196 | 1.93 | Glyma15g18180.1 | AT5G57110.2 | Arabidopsis-autoinhibited Ca2+ -ATPase, isoform 8, contains all of the characteristic motifs of Ca2+ -transporting P-type Ca2+ -ATPases and is localized to the plasma membrane. |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.32014.1.S1_at | 0.06092 | 1.84811 | Glyma18g41240.1 | AT3G53940.1 | mitochondrial substrate carrier family protein; FUNCTIONS IN: transporter activity, binding; INVOLVED IN: transport, mitochondrial transport; LOCATED IN: mitochondrial inner membrane, membrane; EXPRESSED IN: 8 plant structures; EXPRESSED DURING: 4 anthesis, F mature embryo stage, petal differentiation and expansion stage, D bilateral stage, E expanded cotyledon stage; CONTAINS InterPro DOMAIN/s: Mitochondrial carrier protein (InterPro:IPR002067), Mitochondrial substrate carrier (InterPro:IPR001993), Mitochondrial substrate/solute carrier (InterPro:IPR018108), Adenine nucleotide translocator 1 (InterPro:IPR002113); BEST Arabidopsis thaliana protein match is: mitochondrial substrate carrier family protein (TAIR:AT2G37890.1); Has 22773 Blast hits to 10507 proteins in 363 species: Archae - 0; Bacteria - 0; Metazoa - 11593; Fungi - 6281; Plants - 2717; Viruses - 2; Other Eukaryotes - 2180 (source: NCBI BLink). |
| Gma.2198.1.S1_at | 0.07824 | 1.8349 | No Soybean Match Identified, E<10E-30, Perecent ID>95% | AT1G14360.1 | UDP-GALACTOSE TRANSPORTER 3 (UTR3); FUNCTIONS IN: pyrimidine nucleotide sugar transmembrane transporter activity; LOCATED IN: endomembrane system; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: UAA transporter (InterPro:IPR013657); BEST Arabidopsis thaliana protein match is: UTR1 (UDP-GALACTOSE TRANSPORTER 1); UDP-galactose transmembrane transporter/ UDP-glucose transmembrane transporter/ pyrimidine nucleotide sugar transmembrane transporter (TAIR:AT2G02810.1); Has 750 Blast hits to 744 proteins in 166 species: Archae - 0; Bacteria - 0; Metazoa - 423; Fungi - 103; Plants - 108; Viruses - 0; Other Eukaryotes - 116 (source: NCBI BLink). |

| GmaAffx.29926.1.S1_at | 0.05105 | 1.83149 | Glyma06g20360.2 | AT3G47730.1 | member of ATH subfamily |
|---|---|---|---|---|---|
| GmaAffx.25198.1.S1_at | 0.09745 | 1.76468 | No Soybean Match Identified, E<10E-30, Perecent ID>95% | ATMG00640.1 | encodes a plant b subunit of mitochondrial ATP synthase based on structural similarity and the presence in the F(0) complex. |
| GmaAffx.51249.1.S1_at | 0.08371 | 1.75952 | Glyma19g27020.1 | AT1G55620.2 | Encodes a chloride channel protein that has been localized to the chloroplast and golgi apparatus. Complements yeast gef1 mutant and therefor may function to acidify the golgi lumen. |
| GmaAffx.89781.1.S1_s_at | 0.09965 | 1.68881 | Glyma12g00390.2 | AT1G22530.1 | PATELLIN 2 (PATL2); FUNCTIONS IN: transporter activity; INVOLVED IN: transport; LOCATED IN: plasma membrane, chloroplast; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Cellular retinaldehyde-binding/triple function, C-terminal (InterPro:IPR001251), Cellular retinaldehyde-binding/triple function, N-terminal (InterPro:IPR008273), Cellular retinaldehyde binding/alpha-tocopherol transport (InterPro:IPR001071), GOLD (InterPro:IPR009038), Phosphatidylinositol transfer protein-like, N-terminal (InterPro:IPR011074); BEST Arabidopsis thaliana protein match is: PATL1 (PATELLIN 1); transporter (TAIR:AT1G72150.1); Has 52283 Blast hits to 28572 proteins in 1631 species: Archae - 214; Bacteria - 8910; Metazoa - 20172; Fungi - 6486; Plants - 2360; Viruses - 305; Other Eukaryotes - 13836 (source: NCBI BLink). |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.12895.1.A1_at | 0.09251 | 1.62295 | Glyma01g44290.1 | AT5G16880.2 | | VHS domain-containing protein / GAT domain-containing protein; FUNCTIONS IN: protein transporter activity; INVOLVED IN: intracellular protein transport, intra-Golgi vesicle-mediated transport; LOCATED IN: Golgi stack, plasma membrane; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: VHS (InterPro:IPR002014), Target of Myb protein 1 (InterPro:IPR014645), GAT (InterPro:IPR004152), ENTH/VHS (InterPro:IPR008942); BEST Arabidopsis thaliana protein match is: VHS domain-containing protein / GAT domain-containing protein (TAIR:AT1G06210.1); Has 1202 Blast hits to 1198 proteins in 142 species: Archae - 0; Bacteria - 0; Metazoa - 712; Fungi - 289; Plants - 145; Viruses - 0; Other Eukaryotes - 56 (source: NCBI BLink). |
| Gma.13027.1.A1_at | 0.09295 | 1.61331 | Glyma11g35080.1 | AT2G38120.1 | | Encodes an auxin influx transporter. AUX1 resides at the apical plasma membrane of protophloem cells and at highly dynamic subpopulations of Golgi apparatus and endosomes in all cell types. AUX1 action in the lateral root cap and/or epidermal cells influences lateral root initiation and positioning. |
| Gma.259.1.S1_at | 0.05394 | 1.55529 | Glyma05g10930.1 | AT5G55000.2 | | FH protein interacting protein FIP2 |
| Gma.16120.1.A1_x_at | 0.0698 | 1.52769 | Glyma08g23760.1 | AT3G21180.1 | | one of the type IIB calcium pump isoforms. encodes an autoinhibited Ca(2+)-ATPase that contains an N-terminal calmodulin binding autoinhibitory domain. |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.551.1.S1_at | 0.06315 | 1.51468 | Glyma04g41230.1 | AT5G53480.1 | importin beta-2, putative; FUNCTIONS IN: protein transporter activity, binding; INVOLVED IN: intracellular protein transport, protein import into nucleus, docking; LOCATED IN: nucleus, chloroplast, nuclear pore, cytoplasm; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: HEAT (InterPro:IPR000357), Importin-beta, N-terminal (InterPro:IPR001494), Armadillo-like helical (InterPro:IPR011989), Armadillo-type fold (InterPro:IPR016024); BEST Arabidopsis thaliana protein match is: binding / protein transporter (TAIR:AT3G08947.1); Has 1271 Blast hits to 1245 proteins in 165 species: Archae - 0; Bacteria - 0; Metazoa - 604; Fungi - 380; Plants - 120; Viruses - 0; Other Eukaryotes - 167 (source: NCBI BLink). |
| GmaAffx.25445.1.S1_at | 0.06494 | 1.49927 | Glyma09g40760.2 | AT5G01990.1 | auxin efflux carrier family protein; FUNCTIONS IN: auxin:hydrogen symporter activity; INVOLVED IN: auxin polar transport; LOCATED IN: integral to membrane; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Auxin efflux carrier (InterPro:IPR004776); BEST Arabidopsis thaliana protein match is: auxin efflux carrier family protein (TAIR:AT1G71090.1); Has 303 Blast hits to 283 proteins in 69 species: Archae - 0; Bacteria - 2; Metazoa - 0; Fungi - 152; Plants - 93; Viruses - 0; Other Eukaryotes - 56 (source: NCBI BLink). |
| Gma.3339.1.S1_at | 0.06127 | 1.44325 | Glyma15g22520.1 | AT1G30450.2 | |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.90036.1.S1_s_at | 0.02679 | 1.43805 | Glyma08g26150.3 | AT1G72160.1 | SEC14-cytosolic factor family protein / phosphoglyceride transfer family protein; FUNCTIONS IN: transporter activity; INVOLVED IN: transport; LOCATED IN: plasma membrane; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Cellular retinaldehyde-binding/triple function, C-terminal (InterPro:IPR001251), Cellular retinaldehyde-binding/triple function, N-terminal (InterPro:IPR008273), GOLD (InterPro:IPR009038), Phosphatidylinositol transfer protein-like, N-terminal (InterPro:IPR011074); BEST Arabidopsis thaliana protein match is: SEC14 cytosolic factor family protein / phosphoglyceride transfer family protein (TAIR:AT4G09160.1); Has 3452 Blast hits to 2934 proteins in 279 species: Archae - 35; Bacteria - 208; Metazoa - 1309; Fungi - 627; Plants - 480; Viruses - 12; Other Eukaryotes - 781 (source: NCBI BLink). |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.11897.1.S1_at | 0.09726 | 1.41913 | Glyma13g25620.1 | AT5G13750.1 | ZINC INDUCED FACILITATOR-like 1 (ZIFL1); FUNCTIONS IN: tetracycline:hydrogen antiporter activity; INVOLVED IN: response to antibiotic, tetracycline transport; LOCATED IN: integral to membrane; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Tetracycline resistance protein, TetA (InterPro:IPR001958), Major facilitator superfamily MFS-1 (InterPro:IPR011701), Major facilitator superfamily, general substrate transporter (InterPro:IPR016196); BEST Arabidopsis thaliana protein match is: ZIF1 (ZINC INDUCED FACILITATOR 1); carbohydrate transmembrane transporter/ sugar:hydrogen symporter/ tetracycline:hydrogen antiporter (TAIR:AT5G13740.1); Has 10811 Blast hits to 10619 proteins in 1540 species: Archae - 239; Bacteria - 7142; Metazoa - 566; Fungi - 1377; Plants - 204; Viruses - 0; Other Eukaryotes - 1283 (source: NCBI BLink). |
| Gma.8102.1.S1_at | 0.08118 | 1.37642 | Glyma14g08690.1 | AT2G26510.1 | permease |
| GmaAffx.35464.1.A1_at | 0.07249 | 1.36499 | Glyma13g31880.1 | AT1G47670.1 | amino acid transporter family protein; FUNCTIONS IN: amino acid transmembrane transporter activity; INVOLVED IN: amino acid transport; LOCATED IN: plasma membrane, membrane; EXPRESSED IN: 20 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Amino acid transporter, transmembrane (InterPro:IPR013057); BEST Arabidopsis thaliana protein match is: LHT7 (Lys/His transporter 7); amino acid transmembrane transporter (TAIR:AT4G35180.1); Has 952 Blast hits to 951 proteins in 131 species: Archae - 0; Bacteria - 10; Metazoa - 163; Fungi - 114; Plants - 590; Viruses - 0; Other Eukaryotes - 75 (source: NCBI BLink). |
| Gma.3858.2.S1_s_at | 0.0963 | 1.31995 | Glyma08g05940.1 | AT1G67940.1 | member of NAP subfamily |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.32606.1.S1_at | 0.09529 | -3.96792 | Glyma06g46340.1 | AT3G04090.1 | Belongs to a family of plant aquaporins. Similar to yeast and radish aquaporins. Located on ER. |
| GmaAffx.5538.1.A1_at | 0.05603 | -3.24368 | Glyma06g12270.1 | AT1G77380.1 | Amino acid permease which transports basic amino acids. |
| GmaAffx.36217.1.S1_at | 0.07997 | -3.09642 | Glyma09g29410.1 | AT1G48370.1 | Arabidopsis thaliana metal-nicotianamine transporter YSL4 |
| Gma.10462.1.S1_at | 0.03154 | -3.04725 | Glyma07g02680.1 | AT1G15960.1 | member of Nramp2 family |
| GmaAffx.86767.1.S1_at | 0.05952 | -2.98177 | Glyma11g09560.1 | AT1G31770.1 | ABC transporter family protein; FUNCTIONS IN: ATPase activity, coupled to transmembrane movement of substances; INVOLVED IN: transport; LOCATED IN: membrane; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: ATPase, AAA+ type, core (InterPro:IPR003593), ABC transporter-like (InterPro:IPR003439), ABC-2 (InterPro:IPR000412), ABC-2 type transporter (InterPro:IPR013525), ABC transporter, conserved site (InterPro:IPR017871); BEST Arabidopsis thaliana protein match is: ABC transporter family protein (TAIR:AT4G27420.1); Has 231605 Blast hits to 212262 proteins in 2642 species: Archae - 4250; Bacteria - 159285; Metazoa - 7487; Fungi - 4430; Plants - 2737; Viruses - 7; Other Eukaryotes - 53409 (source: NCBI BLink). |
| Gma.2372.1.S1_at | 0.0481 | -2.9013 | No Soybean Match Identified, E<10E-30, Perecent ID>95% | AT2G38940.1 | Encodes Pht1;4, a member of the Pht1 family of phosphate transporters which include: Pht1;1/At5g43350, Pht1;2/At5g43370, Pht1;3/At5g43360, Pht1;4/At2g38940, Pht1;5/At2g32830, Pht1;6/At5g43340, Pht1;7/At3g54700, Pht1;8/At1g20860, Pht1;9/At1g76430 (Plant Journal 2002, 31:341). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.31568.1.S1_at | 0.00857 | -2.74311 | Glyma13g23080.2 | AT5G33280.1 | chloride channel-like (CLC) protein, putative; FUNCTIONS IN: anion channel activity, voltage-gated chloride channel activity; INVOLVED IN: chloride transport; LOCATED IN: intracellular, membrane; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Chloride channel, core (InterPro:IPR014743), Chloride channel, voltage gated (InterPro:IPR001807), Chloride channel plant CLC (InterPro:IPR002251), Cystathionine beta-synthase, core (InterPro:IPR000644); BEST Arabidopsis thaliana protein match is: CLC-C (CHLORIDE CHANNEL C); anion channel/ voltage-gated chloride channel (TAIR:AT5G49890.1); Has 4315 Blast hits to 4052 proteins in 960 species: Archae - 71; Bacteria - 2408; Metazoa - 873; Fungi - 219; Plants - 177; Viruses - 0; Other Eukaryotes - 567 (source: NCBI BLink). |
| GmaAffx.10989.1.S1_at | 0.0931 | -2.74018 | Glyma10g00950.1 | AT2G38940.1 | Encodes Pht1;4, a member of the Pht1 family of phosphate transporters which include: Pht1;1/At5g43350, Pht1;2/At5g43370, Pht1;3/At5g43360, Pht1;4/At2g38940, Pht1;5/At2g32830, Pht1;6/At5g43340, Pht1;7/At3g54700, Pht1;8/At1g20860, Pht1;9/At1g76430 (Plant Journal 2002, 31:341). |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.12965.1.A1_at | 0.05006 | -2.64512 | Glyma06g08910.2 | AT2G25810.1 | | tonoplast intrinsic protein 4;1 (TIP4;1); FUNCTIONS IN: water channel activity; INVOLVED IN: transport; LOCATED IN: integral to membrane, membrane; EXPRESSED IN: 11 plant structures; EXPRESSED DURING: 4 anthesis, F mature embryo stage, petal differentiation and expansion stage, D bilateral stage, E expanded cotyledon stage; CONTAINS InterPro DOMAIN/s: Aquaporin (InterPro:IPR012269), Major intrinsic protein (InterPro:IPR000425); BEST Arabidopsis thaliana protein match is: GAMMA-TIP (GAMMA TONOPLAST INTRINSIC PROTEIN); water channel (TAIR:AT2G36830.1); Has 6492 Blast hits to 6478 proteins in 1243 species: Archae - 59; Bacteria - 2501; Metazoa - 1276; Fungi - 254; Plants - 1459; Viruses - 0; Other Eukaryotes - 943 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.2755.1.A1_at | 0.07811 | -2.42444 | Glyma12g34440.1 | AT3G20660.1 | Arabidopsis thaliana ORGANIC CATION/CARNITINE TRANSPORTER4 (AtOCT4); FUNCTIONS IN: carbohydrate transmembrane transporter activity, sugar:hydrogen symporter activity; INVOLVED IN: transport; LOCATED IN: chloroplast thylakoid membrane, membrane; EXPRESSED IN: 13 plant structures; EXPRESSED DURING: 6 growth stages; CONTAINS InterPro DOMAIN/s: Sugar transporter, conserved site (InterPro:IPR005829), Major facilitator superfamily MFS-1 (InterPro:IPR011701), Major facilitator superfamily, general substrate transporter (InterPro:IPR016196); BEST Arabidopsis thaliana protein match is: AtOCT1 (Arabidopsis thaliana ORGANIC CATION/CARNITINE TRANSPORTER1); carbohydrate transmembrane transporter/ carnitine transporter/ transporter (TAIR:AT1G73220.1); Has 17732 Blast hits to 17562 proteins in 1146 species: Archae - 294; Bacteria - 8210; Metazoa - 4427; Fungi - 3016; Plants - 991; Viruses - 0; Other Eukaryotes - 794 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| Gma.3412.1.S1_at | 0.07148 | -2.18753 | Glyma05g01460.1 | AT5G62670.1 | Arabidopsis H(+)-ATPase 11 (AHA11); FUNCTIONS IN: ATPase activity; INVOLVED IN: cation transport, metabolic process, ATP biosynthetic process; LOCATED IN: plasma membrane, membrane; EXPRESSED IN: 26 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: ATPase, P-type, ATPase-associated region (InterPro:IPR008250), ATPase, P-type cation-transporter, N-terminal (InterPro:IPR004014), Haloacid dehalogenase-like hydrolase (InterPro:IPR005834), ATPase, P-type, H+ transporting proton pump (InterPro:IPR000695), ATPase, P-type, K/Mg/Cd/Cu/Zn/Na/Ca/Na/H-transporter (InterPro:IPR001757), ATPase, P-type, plasma-membrane proton-efflux (InterPro:IPR006534), ATPase, P-type phosphorylation site (InterPro:IPR018303); BEST Arabidopsis thaliana protein match is: AHA4; ATPase/ hydrogen-exporting ATPase, phosphorylative mechanism (TAIR:AT3G47950.1); Has 20808 Blast hits to 18623 proteins in 1829 species: Archae - 405; Bacteria - 11444; Metazoa - 3332; Fungi - 1581; Plants - 1110; Viruses - 3; Other Eukaryotes - 2933 (source: NCBI BLink). |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.4743.1.A1_at | 0.07969 | -2.10932 | Glyma07g17640.1 | | AT3G54140.1 | Encodes a di- and tri-peptide transporter that recognizes a variety of different amino acid combinations. GFP-tagged PTR1 localizes to the plasma membrane and has 8 to 11 predicted transmembrane domains. PTR1 is expressed in a number of different vascular tissues throughout the plant based on promoter:GUS expression analysis. ptr1 mutants have a lower dry weight than wild type plants when both are grown with Pro-Ala or Ala-Ala dipeptides as their nitrogen source, suggesting that PTR1 plays a role in dipeptide uptake in the roots. Furthermore N content of ptr1 mutants is lower than that of wild type plants when grown with Pro-Ala or a mixture of dipeptides as nitrogen source |
| Gma.2250.2.S1_at | 0.03412 | -2.07112 | Glyma20g26500.1 | | AT2G15290.1 | Encodes a protein located in the chloroplast inner envelope. The study of mutant defective in the gene product suggests that the protein is involved in the translocation of protein across the envelope membrane into the chloroplast stroma. |
| GmaAffx.14280.1.S1_at | 0.02316 | -2.05904 | Glyma10g36200.1 | | AT1G22710.1 | Encodes for a high-affinity transporter essential for phloem loading and long-distance transport. A major sucrose transporter, AtSUC2 can also transport a wide range of physiological and synthetic glucose conjugates with both α- or β-linkage. |

| | | | | | | |
|---|---|---|---|---|---|---|
| GmaAffx.51455.1.S1_at | 0.07886 | -2.04667 | Glyma16g31270.3 | AT5G23760.1 | | heavy-metal-associated domain-containing protein; FUNCTIONS IN: metal ion binding; INVOLVED IN: metal ion transport; LOCATED IN: cellular_component unknown; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Heavy metal transport/detoxification protein (InterPro:IPR006121); BEST Arabidopsis thaliana protein match is: heavy-metal-associated domain-containing protein (TAIR:AT1G01490.2); Has 2066 Blast hits to 1059 proteins in 125 species: Archae - 74; Bacteria - 55; Metazoa - 698; Fungi - 142; Plants - 317; Viruses - 16; Other Eukaryotes - 764 (source: NCBI BLink). |
| Gma.528.1.S1_s_at | 0.06518 | -1.92432 | Glyma02g06280.1 | AT1G75220.1 | | integral membrane protein, putative; FUNCTIONS IN: carbohydrate transmembrane transporter activity, sugar:hydrogen symporter activity; INVOLVED IN: transport, transmembrane transport; LOCATED IN: plasma membrane, membrane; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Sugar transporter, conserved site (InterPro:IPR005829), General substrate transporter (InterPro:IPR005828), Sugar/inositol transporter (InterPro:IPR003663), Major facilitator superfamily, general substrate transporter (InterPro:IPR016196); BEST Arabidopsis thaliana protein match is: integral membrane protein, putative / sugar transporter family protein (TAIR:AT1G19450.1); Has 20404 Blast hits to 19903 proteins in 1300 species: Archae - 314; Bacteria - 8577; Metazoa - 4711; Fungi - 4185; Plants - 1437; Viruses - 2; Other Eukaryotes - 1178 (source: NCBI BLink). |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.12855.1.S1_at | 0.06264 | -1.89179 | Glyma05g09150.1 | AT5G14460.1 | | pseudouridine synthase/ transporter; FUNCTIONS IN: pseudouridine synthase activity, transporter activity; INVOLVED IN: tRNA processing, pseudouridine synthesis, RNA modification, tRNA pseudouridine synthesis; LOCATED IN: chloroplast; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: tRNA pseudouridine synthase B, N-terminal, bacterial-type (InterPro:IPR014780), tRNA pseudouridine synthase B, N-terminal (InterPro:IPR002501); BEST Arabidopsis thaliana protein match is: NAP57 (Arabidopsis thaliana homologue of NAP57); pseudouridine synthase (TAIR:AT3G57150.1); Has 5291 Blast hits to 5291 proteins in 1638 species: Archae - 188; Bacteria - 2754; Metazoa - 244; Fungi - 190; Plants - 55; Viruses - 0; Other Eukaryotes - 1860 (source: NCBI BLink). |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.4578.1.S1_s_at | 0.07658 | -1.8897 | No Soybean Match Identified, E<10E-30, Perecent ID>95% | AT2G24520.1 | | Arabidopsis H(+)-ATPase 5 (AHA5); FUNCTIONS IN: ATPase activity; INVOLVED IN: cation transport, metabolic process, ATP biosynthetic process; LOCATED IN: plasma membrane; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: ATPase, P-type, ATPase-associated region (InterPro:IPR008250), ATPase, P-type, H+ transporting proton pump (InterPro:IPR000695), ATPase, P-type cation-transporter, N-terminal (InterPro:IPR004014), Haloacid dehalogenase-like hydrolase (InterPro:IPR005834), ATPase, P-type, K/Mg/Cd/Cu/Zn/Na/Ca/Na/H-transporter (InterPro:IPR001757), ATPase, P-type, plasma-membrane proton-efflux (InterPro:IPR006534), ATPase, P-type phosphorylation site (InterPro:IPR018303); BEST Arabidopsis thaliana protein match is: AHA1 (ARABIDOPSIS H+ ATPASE 1); ATPase/ hydrogen-exporting ATPase, phosphorylative mechanism / protein binding (TAIR:AT2G18960.1); Has 20978 Blast hits to 18423 proteins in 1820 species: Archae - 405; Bacteria - 11682; Metazoa - 3291; Fungi - 1604; Plants - 1083; Viruses - 3; Other Eukaryotes - 2910 (source: NCBI BLink). |
| Gma.11183.1.S1_at | 0.04323 | -1.84631 | Glyma13g24360.1 | AT5G46110.1 | | mutant has Altered acclimation responses; Chloroplast Triose Phosphate Translocator |

| | | | | | |
|---|---|---|---|---|---|
| Gma.4014.1.S1_at | 0.05811 | -1.81355 | Glyma04g34370.1 | AT5G62670.1 | Arabidopsis H(+)-ATPase 11 (AHA11); FUNCTIONS IN: ATPase activity; INVOLVED IN: cation transport, metabolic process, ATP biosynthetic process; LOCATED IN: plasma membrane, membrane; EXPRESSED IN: 26 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: ATPase, P-type, ATPase-associated region (InterPro:IPR008250), ATPase, P-type cation-transporter, N-terminal (InterPro:IPR004014), Haloacid dehalogenase-like hydrolase (InterPro:IPR005834), ATPase, P-type, H+ transporting proton pump (InterPro:IPR000695), ATPase, P-type, K/Mg/Cd/Cu/Zn/Na/Ca/Na/H-transporter (InterPro:IPR001757), ATPase, P-type, plasma-membrane proton-efflux (InterPro:IPR006534), ATPase, P-type phosphorylation site (InterPro:IPR018303); BEST Arabidopsis thaliana protein match is: AHA4; ATPase/ hydrogen-exporting ATPase, phosphorylative mechanism (TAIR:AT3G47950.1); Has 20808 Blast hits to 18623 proteins in 1829 species: Archae - 405; Bacteria - 11444; Metazoa - 3332; Fungi - 1581; Plants - 1110; Viruses - 3; Other Eukaryotes - 2933 (source: NCBI BLink). |

| | | | | | |
|---|---|---|---|---|---|
| GmaAffx.59629.1.S1_at | 0.02679 | -1.76998 | Glyma13g36070.1 | AT3G20660.1 | Arabidopsis thaliana ORGANIC CATION/CARNITINE TRANSPORTER4 (AtOCT4); FUNCTIONS IN: carbohydrate transmembrane transporter activity, sugar:hydrogen symporter activity; INVOLVED IN: transport; LOCATED IN: chloroplast thylakoid membrane, membrane; EXPRESSED IN: 13 plant structures; EXPRESSED DURING: 6 growth stages; CONTAINS InterPro DOMAIN/s: Sugar transporter, conserved site (InterPro:IPR005829), Major facilitator superfamily MFS-1 (InterPro:IPR011701), Major facilitator superfamily, general substrate transporter (InterPro:IPR016196); BEST Arabidopsis thaliana protein match is: AtOCT1 (Arabidopsis thaliana ORGANIC CATION/CARNITINE TRANSPORTER1); carbohydrate transmembrane transporter/ carnitine transporter/ transporter (TAIR:AT1G73220.1); Has 17732 Blast hits to 17562 proteins in 1146 species: Archae - 294; Bacteria - 8210; Metazoa - 4427; Fungi - 3016; Plants - 991; Viruses - 0; Other Eukaryotes - 794 (source: NCBI BLink). |
| GmaAffx.89702.1.A1_at | 0.05811 | -1.73709 | Glyma09g06250.1 | AT4G30190.1 | belongs to the P-type ATPase superfamily of cation-transporting ATPases, pumps protons out of the cell, generating a proton gradient that drives the active transport of nutrients by proton symport. has two autoinhibitory regions within the C-terminal dom |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.312.1.S1_at | 0.0819 | -1.71525 | No Soybean Match Identified, E<10E-30, Perecent ID>95% | AT1G80300.1 | NUCLEOTIDE TRANSPORTER 1 (NTT1); FUNCTIONS IN: ATP:ADP antiporter activity; INVOLVED IN: transport; LOCATED IN: membrane, chloroplast envelope; EXPRESSED IN: stem, guard cell, root, callus, leaf; EXPRESSED DURING: seedling growth; CONTAINS InterPro DOMAIN/s: ADP/ATP carrier protein (InterPro:IPR004667); BEST Arabidopsis thaliana protein match is: ATNTT2; ATP:ADP antiporter (TAIR:AT1G15500.1); Has 587 Blast hits to 585 proteins in 128 species: Archae - 0; Bacteria - 365; Metazoa - 2; Fungi - 15; Plants - 64; Viruses - 0; Other Eukaryotes - 141 (source: NCBI BLink). |
| Gma.16802.1.S1_at | 0.09965 | -1.70041 | Glyma11g09230.1 | AT3G51480.1 | member of Putative ligand-gated ion channel subunit family |
| GmaAffx.19430.1.S1_at | 0.08692 | -1.67943 | Glyma08g21730.1 | AT3G16240.1 | Delta tonoplast intrinsic protein, functions as a water channel and ammonium (NH3) transporter. Highly expressed in flower, shoot, and stem. Expression shows diurnal regulation and is induced by ammonium (NH3). Protein localized to vacuolar membrane. |
| GmaAffx.92088.1.S1_at | 0.09965 | -1.64594 | Glyma08g23830.2 | AT1G51650.1 | ATP synthase epsilon chain, mitochondrial; FUNCTIONS IN: hydrogen ion transporting ATP synthase activity, rotational mechanism, proton-transporting ATPase activity, rotational mechanism; INVOLVED IN: ATP biosynthetic process, ATP synthesis coupled proton transport; LOCATED IN: mitochondrion; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: ATPase, F1 complex, epsilon subunit, mitochondrial (InterPro:IPR006721); Has 170 Blast hits to 170 proteins in 64 species: Archae - 0; Bacteria - 0; Metazoa - 112; Fungi - 3; Plants - 40; Viruses - 0; Other Eukaryotes - 15 (source: NCBI BLink). |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gma.10447.1.S1_at | 0.0958 | -1.57742 | Glyma20g30140.1 | AT4G25640.1 | Encodes a multidrug and toxin efflux family transporter. Involved in flavonoid metabolism, affecting Root growth, seed development and germination, and pollen development, release and viability. |
| Gma.7643.2.S1_at | 0.05603 | -1.49448 | Glyma12g08430.1 | AT5G60790.1 | member of GCN subfamily |
| Gma.2250.1.S1_at | 0.05811 | -1.34434 | Glyma10g40810.1 | AT2G15290.1 | Encodes a protein located in the chloroplast inner envelope. The study of mutant defective in the gene product suggests that the protein is involved in the translocation of protein across the envelope membrane into the chloroplast stroma. |
| Gma.6794.1.A1_at | 0.03473 | -1.2666 | Glyma08g18910.1 | AT5G64290.1 | DICARBOXYLATE TRANSPORT 2.1 (DIT2.1); FUNCTIONS IN: oxoglutarate:malate antiporter activity; INVOLVED IN: malate transport, response to nematode; LOCATED IN: chloroplast, membrane, chloroplast envelope; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Sodium/sulphate symporter (InterPro:IPR001898); BEST Arabidopsis thaliana protein match is: DiT2.2 (dicarboxylate transporter 2.2); oxoglutarate:malate antiporter (TAIR:AT5G64280.1); Has 2721 Blast hits to 2720 proteins in 602 species: Archae - 47; Bacteria - 2055; Metazoa - 43; Fungi - 9; Plants - 101; Viruses - 2; Other Eukaryotes - 464 (source: NCBI BLink). |

**Table 12: qPCR Primers Used for Microarray Validation and Time Course-qPCR**

| Affychip ID | qPCR Primer Name | Primer Sequence (5'-3') |
|---|---|---|
| 004 1 0 0_ | 004 1 0 | |
| | 004 1 0 | |
| 2457 0 0_ | 2457 0 | |
| | 2457 0 | |
| 67310 0 0_ | 67310 0 | |
| | 67310 0 | |
| 737 7 0 0_ | 737 7 0 | |
| | 737 7 0 | |
| 78324 0 0_ _ | 78324 0 | |
| | 78324 0 | |
| 0 13 0 0_ | 0 13 0 | |
| | 0 13 0 | |
| 0 685 0 0_ _ | 0 685 0 | |
| | 0 685 0 | |
| 8442 0 0_ | 8442 0 | |
| | 8442 0 | |
| 6415 0 0_ | 6415 0 | |
| | 6415 0 | |
| 1710 0 0_ | 1710 0 | |
| | 1710 0 | |
| 3478 0 0_ _ | 3478 0 | |
| | 3478 0 | |
| 7475 0 0_ | 7475 0 | |
| | 7475 0 | |
| 057 6 0 0_ | 057 6 0 | |
| | 057 6 0 | |
| 6270 0 0_ | 6270 0 | |
| | 6270 0 | |
| 06732 0 0_ | 06732 0 | |
| | 06732 0 | |
| 0312 0 0_ _ | 0312 0 | |
| | 0312 0 | |
| 7748 0 0_ | 7748 0 | |
| | 7748 0 | |
| 0334 0 0_ | 0334 0 | |
| | 0334 0 | |
| 00223 0 0_ _ | 00223 0 | |
| | 00223 0 | |
| 1028 1 0_ _ | 1028 1 | |
| | 1028 1 | |
| 24 3 1 0_ | 24 3 1 | |
| | 24 3 1 | |
| 31 6 0 0_ | 31 6 0 | |
| | 31 6 0 | |

| | |
|---|---|
| 3454 0 0_ _ | 3454 0 |
| | 3454 0 |
| 4172 0 0_ | 4172 0 |
| | 4172 0 |
| 5 51 0 0_ | 5 51 0 |
| | 5 51 0 |
| 485 0 0_ | 485 0 |
| | 485 0 |
| 288 0 0_ _ | 288 0 |
| | 288 0 |
| 10 68 0 0_ | 10 68 0 |
| | 10 68 0 |
| 05502 0 0_ _ | 05502 0 |
| | 05502 0 |
| 04861 0 0_ | 04861 0 |
| | 04861 0 |
| 0 6 0 1 0_ | 0 6 0 1 |
| | 0 6 0 1 |
| 03161 0 0_ | 03161 0 |
| | 03161 0 |
| 63477 0 0_ | 63477 0 |
| | 63477 0 |
| 6 7 0 0_ | 6 7 0 |
| | 6 7 0 |
| 383 0 0_ | 383 0 |
| | 383 0 |
| 355 2 0 0_ | 355 2 0 |
| | 355 2 0 |
| 18818 0 0_ | 18818 0 |
| | 18818 0 |
| 1633 0 0_ | 1633 0 |
| | 1633 0 |
| 00 3 0 0_ | 00 3 0 |
| | 00 3 0 |
| 77071 0 0_ _ | 77071 0 |
| | _ 77071 0 |
| 3 60 0 0_ | 3 60 0 |
| | 3 60 0 |
| 6626 0 0_ | 6626 0 |
| | 6626 0 |
| **Primers for Timecourse-qPCR** | |
| 57387 0 0_ | 4 |
| | 4 |
| 6512 0 0_ | |
| 355 2 0 0_ | |

**Table 13: Gateway Cloning Primers for Promoter-GUS Constructs.**

| AffyChip ID | Primer Name | Primer Sequence (5'-3') |
|---|---|---|
| 04886 0 0_ | 1 04886 | |
| | 1 04886 | |
| 875 0 0_ | 875 | |
| | 875 | |
| 1358 0 0_ | 1358 | |
| | 1358 | |
| 376 0 0_ | 376 | |
| | 376 | |
| 34 25 0 0_ | 34 25 | |
| | 34 25 | |
| 7 160 0 0_ | 7 160 | |
| | 7 160 | |
| 6036 0 0_ | 6036 | |
| | 6036 | |
| 6276 0 0_ | 6276 | |
| | 6276 | |
| 8442 0 0_ | 8442 | |
| | 8442 | |
| 12 | 12 | |
| | 12 | |
| 0 301 0 0_ | 0 301 | |
| | 0 301 | |
| 0 | 0 | |
| 1 | 1 | |

**Promoter-reporter constructs**

**[0094]** The promoter sequences for the genes used in the GUS reporter assays were identified and downloaded from the soybean genome database (Phytozome, www.phytozome.net). Primers (Table 12) were designed to amplify approximately a 2-kb sequence immediately 5' of the ATG start site (Figure 4b). The promoter DNA fragments were PCR-amplified using soybean cv. Williams 82 genomic DNA as a template and cloned into the gateway cloning vector pDONR-Zeo (Invitrogen, Carlsbad, CA, USA). The cloned promoters were sequenced using vector-specific primers and internal sequencing primers. The correct promoter fragments were then gateway cloned into a pYXT1 vector (Xiao *et al.*, 2005) upstream of a β-glucuronidase (*GUS*) gene as a transcriptional fusion. The final plasmids were verified by PCR analysis and used for transformation of *Agrobacterium rhizogenes* (strain K599).

**Hairy root transformation**

**[0095]** Hairy roots transgenic for each promoter-GUS construct were generated using the method described by Wang *et al.* (2007) with the following modifications. The cotyledons were excised from 9-day-old aseptically grown soybean seedlings (NIL-R or cv. Williams 82) and vacuum infiltrated for 20 min with *A. rhizogenes* culture resuspended in ¼ Gamborg's salt solution (Phytotechnology Lab, Shawnee Mission, KS, USA) carrying various reporter constructs. Cotyledons were co-cultivated with *A. rhizogenes* for 3 days. The cotyledons were later placed on MXB medium [1x MS basal nutrient salts (Gibco BRL), 1x Gamborg's vitamins, 3% w/v sucrose, and 0.8% w/v Daishin agar, pH 5.7] supplemented with kanamycin (200 μg per ml) and timentin (238 μg per ml) and incubated in a growth chamber at 26°C set to a long-day photoperiod (16h light/8h dark). Hairy roots that emerged after 14 days were root-tip propagated twice on MXB medium with kanamycin (200μg per ml) and timentin (238μg per ml), after which the roots were transferred to MXB medium with timentin (237μg per ml). Hairy roots at this stage were either used immediately for nematode inoculation experiments or maintained by subculturing for later use.

**Nematode infection of transgenic hairy roots and GUS staining**

**[0096]** Infective second-stage juveniles (J2) were hatched from eggs as described in Wang *et al.* (2007). Nematodes were surface-sterilized with sterilizing solution (0.004% w/v mercuric chloride, 0.004% w/v sodium azide and 0.002 % v/v Triton X-100) for 8 min followed by 5 washes with sterile water and resuspended in 0.1% w/v agarose. Hairy roots (3-4 cm) grown on MXB medium were inoculated ~1 cm above the root tip with 200 ± 25 J2s per root in

a 25-µl volume. The roots were cut and stained for GUS expression at 5 dpi. GUS staining was done according to Jefferson *et al.* (1987). Briefly, hairy roots were cut 1-2 cm above the infection zone and placed in GUS staining solution (100 mM Tris pH 7.0, 50 mM NaCl, 1 mM X-Gluc, 1.5 mM potassium ferricyanide pH 7.0, 0.06% v/v Triton X-100). The root tissues were vacuum-infiltrated twice for 10 min each and incubated at 37°C overnight. The GUS staining reaction was stopped by replacing staining solution with 70% v/v ethanol. GUS stained roots were photographed under a Leica MZFLIII stereoscope (Leica Microsystems, Bannockburn, IL) fitted with an Optronics MagnaFire, version 2.0, camera (Optronics, Goleta, CA).

**Sample preparation for time-course qPCR analysis**

**[0097]** Infected root tissues for time course qPCR analysis were prepared as described in Ithal *et al.* (2007a), except that samples were collected at 2, 4, 6, and 8 dpi. Excised root pieces from 12-15 different plants were pooled for each genotype/inoculum combination. Samples were quick frozen in liquid nitrogen and stored at -80oC until RNA isolation. Nematode penetration was verified by staining the nematodes in at least five sample roots for each treatment at 24 hours post-inoculation as described by Ithal *et al.* (2007a). Infected root tissues from three independent biological replicates were prepared.

**RNA isolation and qPCR**

**[0098]** Total RNA was isolated from root tissues using the RNeasy plant miniprep kit (Qiagen, Valencia, CA, USA), according to the manufacturer's instructions. First strand cDNA synthesis was carried out using a Superscript III first strand synthesis kit (Invitrogen, USA), according to the manufacturer's instructions. Real-time qPCR was carried out using an Applied Biosystems 7500 real-time PCR system. Gene-specific primers (Table 13) were designed using the primer express software (Applied Biosystems, CA, USA). All qRT-PCR reactions were carried out in triplicate. PCR was performed using the following cycling parameters: 50°C for 2 min, 95°C for 10 min, and 40 cycles of 95°C 15 s and 60°C for 1 min. The soybean ubiquitin gene (Acc. No D28123) was used as an endogenous control. Expression was quantified using the ΔΔCT method in comparison to the endogenous control. Fold-changes were determined relative to the NIL-R mock-inoculated sample for each time point. There were no significant expression differences between mock-treated NIL-S and NIL-R roots.

**Response of NILs to SCN**

**[0099]** Soybean NILs, derived from a cross between the susceptible cultivar Evans and the resistant PI 209332, were chosen for these studies. These NILs are predicted to share 98% of their genome, differing at the major SCN resistance locus, *Rhg1* (Mudge, 1999). NIL-S

is susceptible and NIL-R is resistant to SCN inbred line PA3 (HG type 0). The *Rhg1* allele in PI 209332 is likely similar to the *Rhg1* allele in PI 88788, the source of SCN resistance found in greater than 90% of commercially available SCN-resistant soybean cultivars. Field populations of SCN that can break PI 88788 resistance typically can break PI 209332 resistance, suggesting that these PIs share a similar type of resistance (Colgrove and Niblack, 2008). The delayed resistance response in PI 209332 and PI 88788 is thought to be due to the absence of the *Rhg4* resistance allele which is present in Peking, a cultivar that exhibits a rapid resistance response to SCN. The experimental system disclosed herein takes advantage of this slow resistance response to characterize *Rhg1*-mediated differences in gene expression during syncytium formation.

**[0100]** In laboratory inoculation assays, the penetration and development of SCN and the formation of syncytia on the two NILs were examined to identify appropriate time points for LCM analysis. Freshly hatched second-stage juveniles (J2s) were used for synchronized infection of soybean roots. Roots of two-day-old seedlings were infected with an equal number of SCN (PA3) juveniles, and the roots were acid fuschin-stained at different days post-inoculation (dpi). Infected roots were harvested at different time points and stained with acid fuchsin to monitor the infection process. The results are shown in Fig. 1. Figures 1(a)-(d) show penetration and development of soybean cyst nematode (SCN) PA3 on resistant (NIL-R) and susceptible (NIL-S) lines (a) NIL-S, 2 dpi; (b) NIL-S, 10 dpi; (c) NIL-R, 2 dpi; (d) NIL-R, 10 dpi. Figures 1(e)-(h) show developmental differences between PA3-induced syncytia on NIL-S and NIL-R roots (e) NIL-S syncytium at 5 dpi; (f) NIL-S syncytium at 8 dpi; (g) NIL-R syncytium at 5 dpi; and (h) NIL-R syncytium at 8 dpi. N- nematode, Syn- syncytia. Scale bar: (a)-(d) 250 μm; (e)-(h) 50 μm.

**[0101]** Similar numbers of nematodes were observed in both NIL-R and NIL-S at 2 days post inoculation (dpi) (Figures 1a and 1c) indicating that resistance controlled by *Rhg1* does not affect penetration and migration of nematodes in soybean roots. At 10 dpi significant differences in the development of nematodes were observed between NIL-S and NIL-R. Late fourth-stage juveniles (J4) and early adult females were observed in the NIL-S (Figure 1b) by 10 dpi whereas the majority of the nematodes had only advanced to third-stage (J3) and early J4 stages in the NIL-R (Figure 1d). These results are consistent with results reported by Li *et al.*, (2004) in greenhouse bioassays using these NILs.

**[0102]** To assess the differences in syncytium development at a more refined level, infected root samples at 5, 8, and 10 dpi were sectioned for microscopic examination. At 5 dpi, syncytia appeared normal in both the NIL-S (Figure 1e) and NIL-R (Figure 1g). Normal syncytium development was observed at 8 dpi in NIL-S (Figure 1f) whereas degenerating cells

both in and around developing syncytia were observed in NIL-R at 8 dpi (Figure 1h). The majority of the syncytia were degenerated by 10 dpi in NIL-R. Based on these observations, gene expression changes were assayed within syncytia at 5 and 8 dpi in the NILs.

**Transcript profiling of syncytia in NILs**

**[0103]** The GeneChip Soybean Genome Array (Affymetrix), which carries 37,593 probe sets representing 35,611 soybean transcripts, was used to compare the transcriptional profiles of SCN-induced syncytia in NIL-R and NIL-S. The microarray analysis was carried out using cRNA generated from LCM syncytia at 5 and 8 dpi with SCN from either the NIL-S or NIL-R. No significant evidence of interaction was found between NIL and dpi. Thus, the instant studies were focused on the main effects of NIL and the differences between NIL-S and NIL-R that are averaged over 5 and 8 dpi. This comparison of expression profiles between genotypes resulted in the identification of 1,447 differentially expressed probe sets using a false discovery rate (FDR) set at 10%. Of the 1,447 probe sets, 828 were up-regulated, and 619 were down-regulated in the NIL-R compared to the NIL-S (Tables 1-11). The recently released SoyBase annotation (v.2) for the Affymetrix Soybean Genome Array was used to classify these genes into categories (Figure 2). Of the 1,447 probe sets, 355 (24.5%) correspond to genes coding for unknown proteins and/or those with no known homologs in Arabidopsis, with a high confidence value cut-off (E value <10-6). Two other major categories include cellular metabolism genes (306 probe sets; 21.1%) and stress- and defense-related genes (241 probe sets; 16.8%). Additional classifications include (in descending order) cellular signaling, transporters, proteolysis, transcription factors, protein sorting and transport, cell wall-related, and hormone-related genes. Probe sets that do not fit into any of these categories or fall into multiple categories are grouped as "miscellaneous" (223; 15.4%). A number of probe sets corresponding to different *Glycine max* gene models had the same Arabidopsis homologs; this is not surprising given the duplicated nature of the soybean genome (Schlueter *et al.*, 2004; Schlueter *et al.*, 2007). These probe sets may represent homeologous genes with the same function, especially when their expression patterns fall within ± one-fold difference of each other.

**qPCR validation of microarray data**

**[0104]** The microarray data were validated by qPCR analysis of selected genes using RNA isolated from syncytial cells laser microdissected from the roots of NIL-R and NIL-S at 5 dpi. The genes were selected to represent those that were either up- or down-regulated with fold changes ranging from 27.65 fold up-regulation to 17.63 fold down-regulation in the microarray

analysis (Table 14). Of the 42 genes tested, 38 genes (90.5%) showed differential expression in the same direction as that observed in the microarray experiment (Table 14). Only four probe sets, which showed a down-regulation in the microarray (Gma.2139.2.S1_S_at, GmaAffx.78421.1.S1_at, GmaAffx 84808.1.S1_at, and Gma.3504.2.S1_at, Table 14), were slightly up-regulated in the qPCR analysis. Thus, overall the qPCR results agreed with the microarray results.

**Table 14 qPCR Validation of microarray results**

| SEQ ID | FCaAffyChip Probeset ID | Gene Model/EST Sequence | Putative Function | qPCR FC[a] [b]NIL-R/PA3 5dpi | Microarray FC[a] [b]NIL-R/PA3 5dpi | qPCR FC[a] [c]NIL-R/TN19 5dpi |
|---|---|---|---|---|---|---|
| 1 | Gma.10240.1.A1_at | BE057471 | No predicted gene model | -23.44 | -2.05 | -128.82 |
| 2 | Gma.10701.2.S1_at | Glyma05g27030.1 | P450 pseudogene-like, mandelate racemase N-terminal domain | -1.86 | -4.85 | -2.63 |
| 3 | Gma.10796.1.S1_a_at | Glyma15g43040.1 | Cellulose synthase, CEV1-like | 2.95 | 1.34 | 1.15 |
| 4 | Gma.11004.1.S1_at | Glyma03g35920.1 | Harpin-induced family protein (YLS9)/HIN1 family protein1 | 5.49 | 8.09 | 2.45 |
| 5 | Gma.11334.1.S1_a_at | BI967327 | Caffeoyl-CoA 3-O-methyltransferase, putative | 1.82 | -2.76 | -2.63 |
| 6 | Gma.1423.1.S1_a_at | Glyma01g38450.1 | Brassinosteroid signaling positive regulator-related | 2.57 | 1.65 | 1.91 |
| 7 | Gma.14272.1.S1_at | Glyma10g28300.1 | Hydrophobic protein (RCI2B)/low temp and salt responsive protein (LTI6B) | 33.88 | 11.69 | 2.09 |
| 8 | Gma.1445.1.S1_at | Glyma13g44700.1 | Cinnamoyl-CoA reductase, putative | -1.58 | -3.21 | -1.20 |
| 9 | Gma.15972.1.A1_at | Glyma04g17710.1 | Calcium-binding EF hand family protein | 16.60 | 27.65 | not tested |
| 10 | Gma.16613.1.S1_s_at | Glyma10g40400.1 | Zinc finger (C2H2 type) family protein | 9.55 | 4.67 | 2.00 |
| 11 | Gma.16807.1.S1_at | Glyma09g14880.1 | Zinc finger (B-box type) family protein | 4.37 | 2.68 | 2.40 |
| 12 | Gma.17843.1.S1_at | Glyma04g02660.1 | Gibberellin-regulated protein 1 (GASA1)/gibberellin-responsive protein | 116.98 | 2.62 | 2.75 |
| 13 | Gma.2139.2.S1_s_at | Glyma13g41960.1 | PfkB-type carbohydrate kinase family protein | 1.15 | -1.62 | 1.78 |
| 14 | Gma.2821.1.S1_at | Glyma01g42660.1 | Osmotin-like protein (OSM34) | 19.95 | 6.41 | -2.09 |
| 15 | Gma.3504.2.S1_at | Glyma13g00380.1 | WRKY family transcription factor | 1.48 | -2.59 | 1.95 |
| 16 | Gma.3990.1.S1_s_at | Glyma08g14370.1 | Ethylene-responsive calmodulin-binding protein | 2.09 | 1.9 | 7.76 |
| 17 | Gma.4071.1.S1_at | Glyma17g10050.2 | Gibberellin-regulated protein 2 (GASA2) | -2.75 | -5.26 | -6.17 |
| 18 | Gma.4207.1.S1_at | Glyma06g00630.1 | Myb family transcription factor (MYB32) | -1.23 | -2.15 | 1.78 |
| 19 | Gma.4565.1.S1_s_at | Glyma19g01590.2 | Gibberellin-regulated protein 3 (GASA3) | -1.38 | -4.76 | -20.89 |
| 20 | Gma.4589.1.S1_s_at | Glyma11g35800.1 | Senescence associated protein, SAG | 204.07 | 2.46 | 1.45 |
| 21 | Gma.5283.1.S1_at | Glyma06g07300.1 | Plant natriuretic peptide | -2.19 | -17.63 | -724.44 |
| 22 | Gma.596.1.S1_at | Glyma13g34580.4 | 14-3-3 protein (GRF9) | -1.12 | -2.96 | -3.80 |
| 23 | Gma.6062.1.S1_at | Glyma03g32130.1 | Dehydration-responsive protein-related | -1.26 | -2.69 | 1.45 |
| 24 | Gma.7381.1.S1_at | Glyma20g33430.1 | NAC domain containing protein | 2.45 | 1.67 | 2.34 |
| 25 | Gma.7526.1.A1_at | Glyma11g11430.1 | Senescence/dehydration-associated protein-related (ERD7) | 16.22 | 7.82 | 3.98 |
| 26 | Gma.7737.1.S1_at | BG650195MATE | efflux family protein | 467.74 | 14.73 | 10.96 |
| 27 | Gma.8586.1.S1_at | Glyma17g17330.2 | Proline-rich family protein | 6.92 | 3.39 | 6.61 |
| 28 | Gma.8859.1.A1_at | Glyma15g03110.1 | Hydroxyproline-rich glycoprotein family protein | 21.38 | 5.17 | 3.55 |
| 29 | Gma.9553.1.A1_at | Glyma14g06080.1 | DREB subfamily A-2 of ERF/AP2 transcription factor family | 22.91 | 8.08 | 8.13 |
| 30 | GmaAffx.11502.1.S1_at | Glyma16g01640.1 | Pectinesterase family protein | 5.89 | 2.13 | 2.29 |
| 31 | GmaAffx.21079.1.A1_at | Glyma03g32400.1 | Plasmodesmal protein | 2.95 | 2.15 | 2.04 |
| 32 | GmaAffx.2744.1.S1_at | Glyma04g40930.1 | Auxin-responsive family protein | 6.31 | 3.53 | 2.88 |
| 33 | GmaAffx.29929.1.S1_at | Glyma20g29410.1 | DREB subfamily A-1 of ERF/AP2 transcription factor family (CBF3) | 15.49 | 11.25 | 3.63 |
| 34 | GmaAffx.3568.1.S1_at | Glyma02g19870.1 | bZIP transcription factor family protein (bZIP60 ) | 5.89 | 3.21 | 2.57 |
| 35 | GmaAffx.46603.1.S1_at | Glyma19g27260.1 | GRAM domain-containing protein/ABA-responsive protein-related | 15.14 | 5.55 | 2.29 |
| 36 | GmaAffx.494.1.S1_at | Glyma16g28970.2 | Chitinase A (CHIA) | 6.92 | 4.99 | 10.00 |
| 37 | GmaAffx.70008.1.S1_at | BU762337 | Myb family transcription factor (MYB20) | -131.83 | -3.89 | -630.96 |
| 38 | GmaAffx.74588.1.S1_at | Glyma12g34210.1 | Non-race specific disease resistance protein, NDR1-like | 10.23 | 3.81 | 3.24 |
| 39 | GmaAffx.78421.1.S1_at | Glyma08g15650.1 | Pectinesterase family protein | 1.45 | -1.58 | -1.66 |
| 40 | GmaAffx.84808.1.S1_at | Glyma13g20810.2 | Ethylene-insensitive 2 (EIN2) | 1.91 | -1.59 | 1.82 |
| 41 | GmaAffx.88182.1.S1_at | Glyma03g35930.1 | Harpin-induced family protein/HIN1 family protein | 25.70 | 8.42 | 7.08 |
| 42 | GmaAffx.89435.1.A1_s_at | Glyma03g38190.2 | S-adenosylmethionine synthetase 2 (SAM2) | 2.24 | 1.77 | 1.32 |

[a]Fold-change compared to NIL-S/PA3/5 dpi

[b] "+" indicates genes upregulated; "-" indicates genes downregulated; Note that the fold-changes in expression for certain genes are opposite in

qPCR and microarray;
[c]All of the genes are suppressed by the virulent TN19 SCN population except for those whose fold changes are not tested or underlined; Underlining indicates that those genes are not suppressed by the virulent TN19 SCN population.

**[0105]** A comparative qPCR analysis for these genes was also carried out using RNA isolated from syncytial cells laser microdissected from soybean roots of the NIL-R infected with a virulent SCN population (TN19; HG type 1-7) at 5 dpi. Interestingly, a comparison between qPCR results of syncytia induced in the NIL-R by the virulent and avirulent (PA3; HG type 0) SCN populations showed that the extent of up-regulation or down-regulation of 35 (85.4%) of the 41 genes tested within syncytia induced by the virulent SCN was less than that attained by the avirulent population (Last column, Table 14). These data indicate that the expression of many of these genes is partially suppressed by the virulent SCN population for successful nematode reproduction and development on the NIL-R.

**qPCR analysis of infected whole root pieces**

**[0106]** For purposes of microarray validation and to obtain a detailed temporal expression pattern of select differentially expressed genes, a qPCR analysis was conducted for three genes using RNA isolated from excised SCN-infected whole root pieces at different time points post inoculation. Fig. 3 shows the results of the qPCR analysis of up-regulated genes in excised infected whole root pieces of resistant (NIL-R) and susceptible (NIL-S) near-isogenic lines (NIL) at different days post inoculation (dpi) with avirulent (PA3) or virulent (TN19) soybean cyst nematodes (SCN). Comparison of gene expression levels were made to the NIL-R mock-infected roots (taken as one) at the respective time points post-inoculation. (a) Gma.7623.1.A1_at (b) GmaAffx.68498.1.S1_at (c) GmaAffx.46603.1.S1_at. The qPCR results are normalized to a soybean ubiquitin (Accession No D28123) endogenous control. The graph is representative of 3 independent experiments, and the bars represent confidence intervals in technical replicates as described in Wang *et al.* (2007).

**[0107]** The NIL-R and NIL-S were mock-inoculated or infected with either the avirulent (PA3) or virulent (TN19) SCN population, and root pieces were excised from infection sites at 2, 4, 6, and 8 dpi. Total RNA was isolated from the bulked root pieces (n=10) for qPCR analysis. Time courses of expression were conducted for three different genes (represented by probe sets Gma.7623.1.A1_at, 87.9 fold; GmaAffx.68498.1.S1_at, 13.1 fold; and GmaAffx.46603.1.S1_at, 5.6 fold; Tables 1-11). The expression of mock-inoculated NIL-R at each time point was set as 1. At 2 dpi, the expression of all three genes was more or less equal between treatments (Figure 3). However, by 4 dpi differential expression was observed for all

three genes between the NIL-S and NIL-R. Up-regulation is very clear in the case of genes highly up-regulated in the microarray (Figure 3a, b) and barely detectable in the case of GmaAffx.46603.1.S1_at (Figure 3c), which had the lowest fold up-regulation based on microarray analysis. This study illustrates the dilution effect attributed to using infected whole root pieces for gene expression analyses in this pathosystem. For Gma.7623.1.A1_at, up-regulation in the NIL-R peaks at 6 dpi (Figure 3a). The pattern is similar for GmaAffx.46603.1.S1_at (Figure 3c). The expression pattern is slightly different for the gene represented by probe set GmaAffx.68498.1.S1_at in that the maximum up-regulation is observed at 4 dpi. These peaks in expression levels are consistent with the timing of the resistance response observed in syncytia of the NIL-R (Figure 1). The observed trend in expression for each gene over the time course of infection was reproducible in three independent infection experiments; however, the level of up-regulation varied among experiments. This difference may be attributed to the biological variation inherent to nematode infection experiments, where it is impossible to achieve identical rates of infection. Similar to what was observed in qPCR analyses of RNA isolated from syncytia, infection by the virulent and avirulent SCN populations results in differential up-regulation of these genes. For all three genes tested, the level of up-regulation is lower in response to infection by the virulent nematode population (Figure 3a-c). Additionally, in contrast to infection of the NIL-R, infection of the NIL-S with the avirulent population shows only a slight up-regulation for all three genes tested.

**Promoter-GUS expression analysis**

**[0108]** Promoter-GUS fusions were generated to provide further validation of the spatial expression pattern of the differentially expressed genes identified by microarray analysis and to isolate nematode-responsive soybean promoter sequences with high levels of expression within syncytia. For these experiments, primers corresponding to the 5' upstream sequences of 10 genes (Figure 4a) chosen from the differentially expressed microarray data set (Tables 1-11) were designed by using the recently released Williams 82 soybean genome sequence (Schmutz *et al.*, 2010). Promoter fragments were amplified by PCR using Williams 82 genomic DNA as a template and cloned upstream of a β-glucuronidase (*GUS*) reporter gene in the gateway binary vector pYXT1 (Xiao *et al.*, 2005; Figure 4b).

**[0109]** Transgenic soybean hairy roots were generated in the NIL-R soybean background for each reporter construct. As a positive control, the Arabidopsis *WRKY23* promoter (At2g47260) was tested in soybean; At2g47260 is induced within syncytia in Arabidopsis upon infection with the beet cyst nematode, *Heterodera schachtii* (Grunewald *et*

*al.*, 2008). The transgenic hairy roots were infected with the avirulent (PA3) SCN population. The positive control and all 10 promoter-GUS lines show induced *GUS* gene expression at the nematode feeding sites at 5 dpi (Figure 5). Fig. 5 shows the result of promoter-GUS expression in transgenic soybean hairy root lines of resistant near-isogenic lines (NIL-R) infected with PA3 soybean cyst nematodes (SCN). Promoter-GUS constructs representing ten genes up-regulated in NIL-R identified from the microarray analysis and the nematode-inducible AtWRKY23 were infected with SCN and stained for GUS expression at 5 days post-inoculation (dpi). (a) AtWRKY23 (At2g47260); (b) Glyma15g04570.1; (c) Glyma15g06130.1; (d) Glyma14g06080.1; (e) Glyma09g34110.1; (f) Glyma01g42500.2; (g) Glyma01g42440.1; (h) Glyma18g43750.1; (i) Glyma18g18060.1; (j) Glyma19g04410.1; (k) Glyma13g35100.1. Pictures are representative of at least five independent hairy root lines for each promoter-GUS fusion construct. Syn - syncytia, N - nematode. Scale bar = 500 μm.

**[0110]** For several promoter-GUS lines, GUS expression is observed throughout the root (Figure 4c, d, i, and j) but further induced at nematode feeding sites. Several promoters have an expression pattern that is very low or more restricted to specific cell types within roots and an up-regulation of GUS expression is clearly distinguishable at the nematode feeding sites (Figure 4a, b, e, f, g, h, and k). Figures 6-7 show similar results using two other promoters, namely, Glyma03g35930.1 (88182p) and Glyma03g35920.1(11004p).

**[0111]** Fig. 8 shows longitudinal cross sections of promoter-GUS stained transgenic soybean hairy root lines in the resistant (NIL-R) background infected with PA3 soybean cyst nematodes (SCN). Promoter-GUS stained root pieces of 4 different promoters shown in Figure 5 were fixed in 4% paraformaldehydein phosphate-buffered saline overnight and then paraffin embedded. Serial sections of 10μm thickness were taken. (A) Glyma15g04570.1; (B) Glyma15g06130.1; (C) Glyma14g06080.1; (D) Glyma18g43750.1. Pictures are representative of several syncytiafor each promoter-GUS fusion construct. Syn-syncytia, N -nematode. Scale bar = 50 μm.

**Table 15. Differential expression of stress- and defense-related genes**

| | AffyChip Probeset ID | Q-value | Fold-Change | Gene Model | AT # | Description |
|---|---|---|---|---|---|---|
| 43 | Gma.7623.1.A1_at | 0.0313 | 87.891 | Glyma07g06750.1 | AT2G46240.1 | Similar to AtBAG6 |
| 44 | GmaAffx.75438.1.S1_at | 0.0641 | 62.307 | Glyma17g13820.1 | AT1G76650.1 | Similar to Arabidopsis CML38 |
| 45 | GmaAffx.25859.1.S1_at | 0.0764 | 30.214 | Glyma14g06910.1 | AT3G46230.1 | Class I small heat-shock protein (sHSP) family |
| 46 | GmaAffx.39393.1.S1_at | 0.0347 | 16.162 | Glyma10g32000.1 | AT4G10250.1 | Endomembrane-localized small heat shock protein |
| 47 | GmaAffx.91194.1.S1_at | 0.0376 | 14.499 | Glyma01g41920.1 | AT4G17250.1 | Putative L-lactate/malate dehydrogenase |
| 48 | Gma.17947.1.S1_at | 0.0565 | 13.481 | Glyma08g07340.1 | AT2G29500.1 | Small heat shock protein (HSP17.6B-Class I) |
| 49 | Gma.14272.1.S1_at | 0.0396 | 11.694 | Glyma10g28300.1 | AT3G05890.1 | Rare-cold-inducible 2B (RCI2B) |
| 50 | Gma.10763.1.S1_at | 0.0664 | 11.364 | Glyma04g05720.1 | AT5G12020.1 | Class II heat shock protein (HSP17.6-Class II) |
| 51 | GmaAffx.29929.1.S1_at | 0.0258 | 11.25 | Glyma20g29410.1 | AT4G25480.1 | DREB subfamily A-1 of ERF/AP2 transcription factor family (CBF3) |
| 52 | Gma.2044.1.S1_at | 0.0722 | 10.771 | Glyma09g31740.1 | AT5G66400.2 | ABA and drought-induced glycine-rich dehydrin protein |
| 53 | GmaAffx.68621.1.A1_at | 0.0988 | 9.439 | Glyma20g29770.1 | AT5G52300.2 | Induced in response to water deprivation |
| 54 | GmaAffx.88182.1.S1_at | 0.0448 | 8.423 | Glyma03g35930.1 | AT2G35980.1 | NDR1 and Harpin-like (NHL) gene |
| 55 | Gma.11004.1.S1_at | 0.036 | 8.094 | Glyma03g35920.1 | AT2G35980.1 | NDR1 and Harpin-like (NHL) gene |
| 56 | Gma.9553.1.A1_at | 0.0317 | 8.079 | Glyma14g06080.1 | AT2G40340.1 | DREB subfamily A-2 of ERF/AP2 transcription factor family |
| 57 | Gma.5637.1.S1_at | 0.0595 | 7.885 | Glyma10g28610.1 | AT2G47180.1 | Arabidopsis galactinol synthase 1 (AtGolS1) |
| 58 | Gma.7526.1.A1_at | 0.0312 | 7.621 | Glyma11g11430.1 | AT2G17840.1 | Drought-inducible gene |
| 59 | GmaAffx.63418.1.S1_at | 0.0568 | 7.081 | Glyma09g16810.1 | AT5G06720.1 | Endomembrane located putative peroxidase |
| 60 | Gma.2821.1.S1_at | 0.0362 | 6.407 | Glyma01g42660.1 | AT4G11650.1 | Osmotin-like protein, soybean PR5 |
| 61 | GmaAffx.30428.1.S1_at | 0.0996 | 6.375 | Glyma17g08020.1 | AT1G16030.1 | Heat shock protein 70B (HSP70b) |
| 62 | Gma.8386.1.S1_at | 0.0258 | 6.08 | Glyma04g37040.1 | AT1G76650.1 | Similar to Arabidopsis CML38 |
| 63 | GmaAffx.6438.1.S1_at | 0.0678 | 5.926 | Glyma11g29720.1 | AT2G38470.1 | Similar to WRKY33 |
| 64 | GmaAffx.92590.1.S1_at | 0.0766 | 5.848 | Glyma02g33070.1 | AT2G41680.1 | NADPH thioredoxin sulphide reductase |
| 65 | Gma.13195.1.S1_s_at | 0.0274 | 5.445 | Glyma01g20860.1 | AT1G13340.1 | Unknown protein: oxidative stress |
| 66 | GmaAffx.6519.1.S1_at | 0.0624 | 5.408 | Glyma08g29470.1 | AT1G13340.1 | Unknown protein: oxidative stress |
| 67 | Gma.1622.1.A1_s_at | 0.0273 | 5.246 | Glyma05g17470.1 | AT5G66900.1 | Putative disease resistance protein (CC-NBS-LRR class) |
| 68 | Gma.5820.1.S1_s_at | 0.0258 | 4.856 | Glyma06g19810.1 | AT1G19020.1 | Unknown protein; oxidative stress |
| 69 | Gma.1439.1.S1_at | 0.0317 | 4.217 | Glyma04g14800.3 | AT3G22370.1 | Alternative oxidase AOX1a |
| 70 | GmaAffx.71308.2.A1_at | 0.0337 | 4.003 | Glyma04g05500.2 | AT2G26150.1 | Similar to AtHSF-A2 |
| 71 | Gma.620.1.S1_at | 0.0661 | 3.931 | Glyma15g40190.1 | AT1G78380.1 | Glutathione transferase of Tau GST gene family |
| 72 | GmaAffx.74588.1.S1_at | 0.0398 | 3.806 | Glyma12g34210.1 | AT3G20600.1 | NDR1-like |
| 73 | Gma.4089.1.S1_at | 0.0276 | 3.722 | No Soybean Match | AT4G25200.1 | AtHSP23.6-mito mRNA |
| 74 | GmaAffx.36259.1.S1_s_at | 0.0581 | 3.493 | Glyma05g33200.1 | AT2G02120.1 | Plant defensin (PDF) family PR protein |
| 75 | Gma.8204.1.A1_at | 0.0771 | 3.462 | Glyma04g14800.3 | AT3G22370.1 | Alternative oxidase AOX1a |
| 76 | Gma.8456.1.S1_at | 0.0922 | 3.265 | Glyma08g20190.1 | AT1G55020.1 | Lipoxygenase, similar to AtLOX1 |
| 77 | GmaAffx.3568.1.S1_at | 0.0196 | 3.208 | Glyma02g19870.1 | AT1G42990.1 | Similar to AtbZIP60 |
| 78 | GmaAffx.92499.1.S1_s_at | 0.0613 | 3.123 | Glyma19g43460.1 | AT3G04720.1 | Similar to antifungal chitin-binding protein hevein, PR-4 |
| 79 | GmaAffx.80951.1.S1_at | 0.0682 | 3.042 | Glyma16g29750.1 | AT5G52640.1 | AtHSP90.1 homolog |
| 80 | GmaAffx.84566.1.S1_x_at | 0.0539 | 3.011 | Glyma18g49360.1 | AT3G28910.1 | Transcription factor AtMYB30 homolog |
| 81 | Gma.7922.1.A1_a_at | 0.0431 | 2.88 | Glyma05g02210.1 | AT5G62520.1 | Similarity to RCD1 but without the WWE domain |
| 82 | GmaAffx.92919.1.S1_at | 0.0878 | 2.674 | Glyma05g05290.1 | AT5G47120.1 | Encodes BI-1, a homolog of mammalian Bax inhibitor 1 |
| 83 | GmaAffx.93596.1.S1_at | 0.040 | 2.672 | Glyma17g03950.2 | AT5G49520.1 | Similar to WRKY48 |
| 84 | GmaAffx.34450.1.S1_at | 0.0734 | 2.669 | Glyma05g05290.1 | AT5G47120.1 | Encodes BI-1, a homolog of mammalian Bax inhibitor 1 |
| 85 | GmaAffx.19934.1.S1_at | 0.0416 | 2.593 | Glyma13g21490.2 | AT5G03720.1 | Heat stress transcription factor HSFA3 |
| 86 | GmaAffx.1338.1.S1_at | 0.0962 | 2.555 | Glyma02g35210.1 | AT3G11820.1 | Similar to SYP121(PENETRATION1/PEN1) |
| 87 | Gma.8336.1.S1_at | 0.047 | 2.398 | Glyma19g40560.1 | AT2G47260.1 | Similar to WRKY23 |
| 88 | Gma.4639.1.A1_at | 0.0963 | 2.281 | Glyma13g36340.1 | AT3G20600.1 | NDR1-like |
| 89 | GmaAffx.22821.1.S1_at | 0.0687 | 2.178 | Glyma19g44390.2 | AT3G03300.2 | Encodes a dicer-like 2 |
| 90 | Gma.11115.2.S1_at | 0.0268 | 2.146 | Glyma11g14950.1 | AT3G12580.1 | Heat shock protein 70 (HSP70) |
| 91 | GmaAffx.11781.1.S1_s_at | 0.0273 | 2.137 | Glyma02g35660.1 | AT5G06320.1 | NDR1 and Harpin-like (NHL) gene |
| 92 | GmaAffx.89654.1.A1_s_at | 0.0338 | 2.063 | No Soybean Match | AT1G21750.1 | Protein disulfide isomerase-like (PDIL) protein |
| 93 | GmaAffx.20155.1.S1_at | 0.0517 | 2.036 | Glyma02g35230.1 | AT3G11820.1 | Similar to SYP121(PENETRATION1/PEN1) |
| 94 | Gma.10639.1.S1_x_at | 0.065 | 1.793 | Glyma03g27560.1 | AT2G41010.1 | Similar to a calmodulin binding protein |
| 95 | Gma.6474.1.A1_s_at | 0.0655 | 1.727 | Glyma15g13470.1 | AT2G34690.1 | Similar to ACD11 gene |
| 96 | GmaAffx.1991.1.S1_at | 0.0637 | 1.665 | Glyma11g04040.2 | AT5G47120.1 | Encodes BI-1, a homolog of mammalian Bax inhibitor 1 |
| 97 | GmaAffx.80733.1.S1_at | 0.0396 | 1.578 | Glyma17g11240.1 | AT5G20320.1 | Encodes a dicer-like 4 |
| 98 | Gma.3755.1.S1_at | 0.0869 | 1.481 | Glyma03g30590.1 | AT5G13320.1 | Similar to PBS3/GH3.12 |
| 99 | GmaAffx.19777.1.A1_at | 0.0986 | 1.43 | No Soybean Match | AT4G31800.2 | Similar to WRKY18 pathogen-induced transcription factor |
| 100 | GmaAffx.48022.2.A1_at | 0.0658 | 1.429 | Glyma13g02470.3 | AT4G08500.1 | Encodes MEKK1, phosphorylates MEK1 |

**[0112]** Special attention was paid to the stress- and defense-related genes to gain a better understanding of the HR that occurs within developing syncytia of the NIL-R in response to SCN. The differential expression of this class of genes varies from 87-fold up-regulated to 17-fold down-regulated (Tables 15-16). In one embodiment, the expression of any of the genes

listed in Tables 14-16 may be altered in order to obtain a nematode resistant plant. In another embodiment, the expression levels of up-regulated genes such as those listed in Table 15 may be increased in order to boost plant defense against nematode infection. In another embodiment, the expression levels of down-regulated genes such as those listed in Table 16 may be decreased in a host plant in order to boost plant defense against nematode infection.

**[0113]** Soybean orthologs of many known plant defense genes have not yet been identified; therefore, we relied on their similarity to Arabidopsis homologs. A total of 241 probe sets representing 16.8% of the total number of differentially expressed genes identified are classified in this group. These included genes involved in apoptosis and disease resistance. A large number of genes involved in oxidative, heat, drought, cold, osmotic, and salt stress responses are also differentially regulated. A natriuretic peptide with an expansin-like domain and many abscisic acid (ABA) induced genes are also among those differentially expressed (Tables 15-16).

**[0114]** The gene up-regulated with the highest fold change (87-fold, Table 15) is a probe set that corresponds to a soybean gene with similarity to Arabidopsis Bag6 (*AtBag6*). *Bag6* encodes a stress-induced calmodulin-binding BAG (BCL2-associated athanogene) domain protein that is homologous to mammalian BAG proteins, which are regulators of BCL2 involved in apoptosis (Kang *et al.*, 2006). Overexpression of *Bag6* in yeast and Arabidopsis causes cell death (Kang *et al.*, 2006). Another highly up-regulated gene (62-fold) is a soybean calmodulin with similarity to calmodulin-like 38 (CML38) of Arabidopsis, which has been shown to respond to wounding (Vanderbeld and Snedden, 2007). Several heat shock proteins (HSPs) of the small HSP superfamily are up-regulated as well (Table 15). HSPs are stress responsive proteins that have a protective function in promoting cellular stress tolerance (Wang *et al.*, 2004). Small HSPs bind and stabilize denatured proteins to which other high molecular weight HSPs act as chaperones under stress conditions. Several other HSPs, including an HSP70 homolog (Gma.11115.2.S1_at), HSP70B homolog (GmaAffx.30428.1.S1_at), HSP90.1 homolog (GmaAffx.80951.1.S1_at), and two heat shock transcription factors (HSFs), Hsf-A2 homolog (GmaAffx.71308.2.A1_at, 4.0 fold) and Hsf-A3 homolog (GmaAffx.19934.1.S1_at, 2.6 fold), are up-regulated in syncytia of the NIL-R. HSP90 is a highly conserved molecular chaperone rapidly induced during pathogen challenge and a variety of environmental stresses. It interacts with the R protein, RPM1 (Hubert *et al.*, 2003), and is required for *RPS2*-mediated resistance against *Pseudomonas syringae* pv. tomato DC 3000 (*avrRpt2*) (Takahashi *et al.*, 2003). HSFs are involved in a variety of environmental stresses; HSF-A2, for example, is a key inducer of defense responses and is up-regulated during environmental stress and $H_2O_2$

treatment (Nishizawa *et al.*, 2006). Several PR genes are also up-regulated. A soybean osmotin (Gma.2821.1.S1_at, Table 15), which is described as a salt stress-induced acidic isoform of PR-5 (Onishi *et al.*, 2006) and has similarity to Arabidopsis osmotin 34, is up-regulated 6.4 fold. Osmotins are components of incompatible reactions against bacterial pathogens (Jia and Martin, 1999). Another up-regulated PR-protein is a hevein-like protein belonging to the PR-4 family, which is up-regulated during salt stress, in response to viral infection, and in systemic acquired resistance (SAR) (Potter *et al.*, 1993). We found a 3.5-fold up-regulation of a defensin homologous to Arabidopsis defensin PDF2.1 (GmaAffx.36259.1.S1_s_at) but a down-3.5-fold down-regulation of another member of the same defensin family, PDF2.5 (Gma.4126.1.S1_at, Table 16).

**Table 16 Down-regulated Stress and Defense-related genes**

| | AffyChip Probeset ID | Q-value | Fold-Change | Gene Model | AT # | Description |
|---|---|---|---|---|---|---|
| 101 | Gma.5283.1.S1_at | 0.037 | -17.624 | Glyma06g07300.1 | AT2G18660.1 | Plant natriuretic peptide A (PNP-A) |
| 102 | Gma.5629.2.S1_a_at | 0.081 | -5.461 | Glyma15g05820.1 | AT2G41480.1 | Peroxidase, response to oxidative stress |
| 103 | GmaAffx.84317.1.S1_at | 0.054 | -5.457 | Glyma06g42310.1 | AT5G54250.2 | Similar to AtCNGC4, HR |
| 104 | Gma.5629.1.S1_at | 0.032 | -5.021 | No Soybean Match | AT2G41480.1 | Peroxidase, response to oxidative stress |
| 105 | Gma.5971.1.S1_at | 0.047 | -4.813 | Glyma11g07670.1 | AT5G66390.1 | Similar to Arabidopsis PER72, oxidative stress |
| 106 | GmaAffx.74124.1.S1_at | 0.083 | -3.937 | No Soybean Match | AT5G67400.1 | Similar to PER73(RHS19), oxidative stress |
| 107 | Gma.1539.1.S1_at | 0.064 | -3.677 | Glyma11g05300.2 | AT4G37520.1 | Peroxidase similar to PER50, oxidative stress |
| 108 | GmaAffx.73002.1.S1_at | 0.060 | -3.570 | Glyma02g09470.1 | AT1G14790.1 | Similar to RDRP |
| 109 | Gma.4126.1.S1_at | 0.058 | -3.461 | Glyma06g16810.1 | AT5G63660.1 | Similar to defensin PDF2.5 defense response |
| 110 | Gma.4919.1.S1_at | 0.027 | -3.198 | Glyma14g05840.1 | AT5G05340.1 | Peroxidase, response to oxidative stress |
| 111 | GmaAffx.92030.1.S1_at | 0.095 | -2.594 | Glyma19g44310.1 | AT2G46370.2 | Similar to JAR1 |
| 112 | Gma.3504.2.S1_at | 0.050 | -2.591 | Glyma13g00380.1 | AT4G31550.1 | Similar to WRKY11 |
| 113 | GmaAffx.65280.1.A1_at | 0.078 | -2.574 | Glyma13g26640.2 | AT3G27890.1 | NADPH quinone oxidoreductase |
| 114 | Gma.8020.3.S1_at | 0.074 | -2.521 | Glyma13g07490.1 | AT3G25780.1 | Similar to AOC3 |
| 115 | Gma.338.1.S1_at | 0.027 | -2.438 | Glyma07g39020.1 | AT4G21960.1 | Peroxidase, response to oxidative stress |
| 116 | Gma.1502.1.S1_at | 0.087 | -2.285 | Glyma05g29400.1 | AT2G29420.1 | Similar to Arabidopsis glutathione S-transferase tau 7 |
| 117 | Gma.2350.1.S1_at | 0.054 | -2.227 | No Soybean Match | AT5G60640.1 | PDIL-4 homolog, oxidative stress |
| 118 | Gma.8020.2.S1_a_at | 0.084 | -2.112 | No Soybean Match | AT1G13280.1 | Similar to AOC4, jasmonic acid biosynthesis |
| 119 | Gma.4207.1.S1_at | 0.027 | -2.147 | Glyma06g00630.1 | AT4G34990.1 | AtMYB32 homolog |
| 120 | GmaAffx.6478.1.S1_s_at | 0.068 | -2.051 | Glyma13g00380.1 | AT4G31550.1 | Similar to WRKY11 |
| 121 | Gma.3504.1.S1_at | 0.095 | -1.965 | Glyma17g06450.1 | AT4G31550.1 | Similar to WRKY11 |
| 122 | Gma.4189.1.S1_at | 0.054 | -1.942 | Glyma17g01720.1 | AT4G21960.1 | PRXR1, oxidative stress |
| 123 | Gma.3504.2.S1_a_at | 0.079 | -1.923 | Glyma13g00380.1 | AT4G31550.1 | Similar to WRKY11 |
| 124 | GmaAffx.54278.1.S1_at | 0.078 | -1.782 | Glyma13g03600.1 | AT1G21750.1 | Similar to PDIL1-1, regulation of PCD |
| 125 | Gma.2677.1.S1_s_at | 0.064 | -1.731 | Glyma19g44310.1 | AT2G46370.2 | Similar to At AR1, production of JA-Ile |
| 126 | Gma.2749.1.S1_at | 0.048 | -1.699 | Glyma10g40140.1 | AT1G80600.1 | Similar to WIN1, defense response |
| 127 | Gma.2350.1.S1_s_at | 0.097 | -1.645 | Glyma13g40130.1 | AT5G60640.1 | PDI1-4, oxidative stress |
| 128 | Gma.4312.1.S1_at | 0.098 | -1.633 | Glyma08g05200.1 | AT2G31570.1 | Glutathione peroxidase 2 (GPX2) |
| 129 | GmaAffx.89649.1.S1_s_at | 0.052 | -1.631 | Glyma05g34490.4 | AT2G43350.1 | ATGPX3 |
| 130 | GmaAffx.84808.1.S1_at | 0.039 | -1.590 | Glyma13g20810.2 | AT5G03280.1 | Ethylene insensitive 2 (EIN2) |
| 131 | GmaAffx.85352.1.S1_at | 0.091 | -1.542 | Glyma03g33850.1 | AT5G03280.1 | Ethylene insensitive 2 (EIN2) |
| 132 | Gma.4312.1.S1_x_at | 0.097 | -1.461 | Glyma08g05200.1 | AT2G31570.1 | Glutathione peroxidase 2 (GPX2) |
| 133 | Gma.3301.1.S1_at | 0.093 | -1.426 | Glyma06g00440.1 | AT4G02600.2 | Homology to MLO1 protein, cell death defense response |
| 134 | GmaAffx.90444.1.S1_s_at | 0.058 | -1.323 | Glyma11g21260.1 | AT3G27890.1 | NADPH quinone oxidoreductase |

**[0115]** Several genes related to oxidative stress also were identified as differentially expressed, indicating that developing syncytia at 5 dpi are undergoing severe oxidative stress (Table 15). The production of reactive oxygen species (ROS) is a key aspect of the HR during R-mediated resistance to other pathogens (Lamb and Dixon, 1997). An NADPH thioredoxin

reductase, similar to Arabidopsis NTRC, is up-regulated (5.8 fold, GmaAffx.92590.1.S1_at). Alternative oxidase (Gma.1439.1.S1_at, 4.2 fold; Gma.8204.1.A1_at, 3.46 fold), glutathione S-transferase (Gma.620.1.S1_at, 3.9 fold), and a gene similar to RCD1-5 involved in ROS regulation (Gma.7922.1.A1_a_at, 2.9 fold) are up-regulated. Several genes related to oxidative stress and regulation of ROS are also down-regulated (Table 16) as are many peroxidases (Gma.5629.2.S1_a_at, Gma.5629.1.S1_at, Gma.5971.1.S1_at, GmaAffx.74124.1.S1_at, Gma.1539.1.S1_at, Gma.4919.1.S1_at, Gma.338.1.S1_at, Gma.4189.1.S1_at, fold changes ranging from -5.4 to -1.9). Peroxidases are involved in H2O2 catabolism, and their down-regulation may suggest a positive impact on ROS generation; although, they can also generate ROS species (Passardi *et al.*, 2004). Other down-regulated oxidative stress genes include two NADPH quinone oxidoreductases (GmaAffx.65280.1.A1_at, -2.57 fold; GmaAffx.90444.1.S1_s_at, -1.3 fold), glutathione peroxidase 2 and 3 homologs, and a protein disulphide isomerase-like 4 (PDI like-4) that belongs to the thioredoxin family.

**[0116]** Several soybean genes showing high similarity to defense genes of Arabidopsis that play a role in incompatible responses to other plant pathogens were found to be differentially expressed in syncytia of the NIL-R in response to SCN (Tables 15-16). Two soybean genes with homology to Arabidopsis *NDR1* are up-regulated (GmaAffx.74588.1.S1_at, 3.8 fold; Gma.4639.1.A1_at, 2.3 fold). NDR1 is involved in SA-mediated disease resistance to biotrophic pathogens (Century *et al.*, 1995). In addition, several *NDR1* and *harpin-like* (*NHL*) genes are up-regulated (GmaAffx.88182.1.S1_at, 8.4 fold; Gma.11004.1.S1_at, 8.1 fold; GmaAffx.11781.1.S1_s_at, 2.1 fold; Table 15). *NHL3* and *NHL10* are induced in response to avirulent viral infection, in senescing leaves, and by spermine in Arabidopsis (Zheng *et al.*, 2004). WRKY transcription factors are known to take part in defense responses to viral, bacterial, and fungal pathogens (Eulgem and Somssich, 2007). Several WRKY transcription factor homologs are up-regulated in syncytia of the NIL-R (Table 15). Also up-regulated are a homolog of AtWRKY33, a known regulator of defense pathways mediating resistance to *P. syringae* and fungal necrotrophic pathogens (Zheng *et al.*, 2006) (GmaAffx.6438.1.S1_at, 5.9 fold); an AtWRKY48 homolog (GmaAffx.93596.1.S1_at, 2.67 fold); and a homolog of AtWRKY18, which is involved in SA-mediated defenses against viruses, bacteria, and fungi (GmaAffx.19777.1.A1_at, 1.4 fold). Interestingly, a soybean homolog of AtWRKY23 (Gma.8336.1.S1_at, 2.4 fold), which is involved in nematode feeding site establishment (Grunewald *et al.*, 2008), is also up-regulated. Down-regulated WRKYs (Table 16) include a homolog of AtWRKY11 (GmaAffx.6478.1.S1_s_at, -2.05 fold; Gma.3504.1.S1_at, -2 fold; Gma.3504.2.S1_a_at, -1.9 fold), a negative regulator of basal defense responses against bacterial

pathogens (Journot-Catalino *et al.*, 2006). Down-regulation of a negative regulator would lead to an enhanced defense response.

**[0117]** Several genes known to take part in SA-mediated defense responses were found to be either up-regulated or down-regulated. Probe set GmaAffx.84566.1.S1_x_at (Table 15) is up-regulated 3-fold; this probe set corresponds to a soybean MYB protein homologous to AtMYB30, an SA-dependent R2-R3 MYB that acts as a positive regulator of HR cell death and is a modulator of SA levels (Vailleau *et al.*, 2002; Raffaele *et al.*, 2006). Soybean homologs of Arabidopsis ACD11 (Gma.6474.1.A1_s_at, 1.7 fold) and PBS3 (Gma.3755.1.S1_at, 1.5 fold), which are involved in SA-mediated defense, are also up-regulated. A soybean homolog of Arabidopsis WIN1 (Gma.2749.1.S1_at, -1.7 fold), a negative regulator of SA accumulation, is down-regulated (Table 16).

**[0118]** Jasmonate-mediated response components are also differentially expressed. A soybean *LOX* gene homologous to Arabidopsis *LOX1* is up-regulated (Gma.8458.1.S1_at, 3.3 fold; Table 15). However, a JAR1 homolog (GmaAffx.92030.1.S1_at, -2.6 fold; Gma.2677.1.S1_s_at, -1.7 fold) and AOC homologs (Gma.8020.3.S1_at AOC3, -2.5 fold; Gma.8020.2.S1_a_at, -2.1 fold; Table 16) are down-regulated.

**[0119]** Also up-regulated are a CC-NB-LRR protein (Gma.1622.1.A1_s_at, 5.2 fold); genes associated with endoplasmic reticulum (ER) stress and the unfolded protein response (UPR), including an AtbZIP60 homolog (GmaAffx.3568.1.S1_at, 3.2 fold) and an AtBIP2 homolog (Gma.17631.1.S1_at, 2 fold) (Table 15); and a Bax inhibitor protein (BAX-I) homolog, an attenuator of apoptosis, (GmaAffx.1991.1.S1_at; GmaAffx.34450.1.S1_at; GmaAffx.92919.1.S1_at, 2.7-1.7 fold). The up-regulation of a MAP3K homolog (GmaAffx.48022.2.A1_at) implicates MAPK signaling in the regulation of resistance to SCN. A soybean gene encoding a protein with homology to Arabidopsis Syntaxin 121 (SYP121), a secretory pathway protein with known roles in defense responses, is also induced (GmaAffx.1338.1.S1_at, 2.6-fold GmaAffx.20155.1.S1_at 2-fold) as are several genes involved in cold, drought, dehydration, and ABA responses (Gma.14272.1.S1_at, 11.7 fold; Gma.2044.1.S1_at, 10.8 fold; GmaAffx.68621.1.A1_at, 9.4 fold; Gma.7526.1.A1_at, 7.6 fold) and two transcription factors of the AP2/ERF family involved in drought responses (GmaAffx.29929.1.S1_at, 11.2 fold; Gma.9553.1.A1_at, 8.1 fold). The high up-regulation of these genes may suggest new roles in HR against a biotrophic pathogen, or they are secondary physiological responses that potentiate HR.

**[0120]** The most highly down-regulated probe set (Gma.5283.1.S1_at, -17.6 fold) corresponds to a gene encoding a predicted natriuretic peptide with an expansin-like domain

sharing homology to AtPNP-A (Table 16), which is involved in plant growth and homeostasis (Morse *et al.*, 2004). AtPNP-A is induced by SA and is expressed at higher levels in Arabidopsis mutants with increased SA levels (Meier *et al.*, 2008). Another highly down-regulated gene is a cyclic nucleotide-gated channel (CNGC) (GmaAffx.84317.1.S1_at, -5.5 fold, Table 16), which shares homology with Arabidopsis CNGC4/HLM1. Arabidopsis mutants of CNGC4/HLM1 produce a lesion mimic phenotype and altered HR (Balague *et al.*, 2003).

**[0121]** Besides the defense-related transcription factors such as WRKYs and AP2/ERF, several other classes of transcription factors were found to be differentially expressed (Table 17). Major classes represented were the NAC domain transcription factors, C2H2-type zinc finger transcription factors, transcription factors involved in cell fate determination, and auxin response factors (ARFs). EIN3 (GmaAffx.65341.1.A1_at, 10 fold) which acts downstream of the histidine kinase ethylene receptor, ETR1 to regulate the ethylene signaling pathway (Chao et al., 1997) was upregulated. Several NAC domain factors were up-regulated. Two probe sets corresponding to soybean genes homologous to ANAC073 (Gma.8113.1.A1_at, 15 fold; GmaAffx.50811.2.S1_at, 11 fold) were upregulated. ANAC073 is involved in the regulation of secondary wall thickening (Zhong et al., 2008). Other NACs included an ANAC087 homolog (GmaAffx.64710.1.S1_at, 4.6 fold; GmaAffx.5448.1.S1_at, 2.5 fold), an ANAC014 homolog (GmaAffx.9475.1.A1_s_at, 1.84 fold), ATAF1 (GmaAffx.90028.1.S1_s_at, 1.9 fold), and an ANAC050 homolog (Gma.7381.1.S1_at, 1.7 fold). ATAF1 has been shown to be induced by wounding and is a negative regulator of defense to bacterial and fungal pathogens (Wang et al., 2009). Four C2H2-type zinc finger transcription factors were up-regulated (Gma.986.1.S1_at, 11 fold; Gma.17736.1.S1_at, 7.6 fold; Gma.4526.1.S1_at, 4.1 fold; GmaAffx.65885.1.A1_s_at, 1.9 fold). The role of these genes in defense or pathogenesis is currently unknown. A homolog of ZAT11 (Gma.4526.1.S1_at) is known to be up-regulated by $H_2O_2$ (Gechev et al., 2005). A MYB domain 20 homolog (GmaAffx.70008.1.S1_at, -3.9 fold) involved in secondary wall thickening was down-regulated. ARF family genes play a central role in controlling sensitivity to the plant hormone auxin. A soybean homolog of ARF19 implicated in root cap development was up-regulated (Gma.9082.1.S1_at, 1.5 fold). Soybean homologs of ARF16 and ARF8 were down-regulated. In Arabidopsis, ARF16 is indispensible for root cap development and is regulated by miR160 (Wang et al., 2005). ARF8, along with ARF6 has been shown to affect jasmonate production in flowers (Nagpal et al., 2005).

Table 17 Differentially Expressed Transcription Factors.

| SEQ ID | AffyChip Probeset ID | Q-value | Fold-Change | Gene Model | AT # | Description |
|---|---|---|---|---|---|---|
| 135 | GmaAffx.65341.1.A1_at | 0.079 | 10.012 | Glyma08g14630.1 | AT5G10120.1 | EIN3 family proteln |
| 136 | Gma.8113.1.A1_at | 0.027 | 15.780 | Glyma15g08480.2 | AT4G28500.1 | NAC domain similar to ANAC073 |
| 137 | GmaAffx.50811.2.S1_at | 0.058 | 11.475 | Glyma15g08480.2 | AT4G28500.1 | NAC domain similar to ANAC073 |
| 138 | Gma.986.1.S1_at | 0.034 | 11.089 | Glyma15g04570.1 | AT2G28710.1 | Zinc finger (C2H2) family protein |
| 139 | Gma.17736.1.S1_at | 0.054 | 7.533 | Glyma03g33070.1 | AT2G37430.1 | Zinc finger (C2H2) family protein (ZAT11) |
| 140 | GmaAffx.41136.1.S1_at | 0.063 | 5.735 | No Soybean Match Identified, E<10E-30, Perecent ID>95% | AT2G34710.1 | Similar to PHABULOSA, leaf pattern formation |
| 141 | Gma.16700.1.S1_at | 0.067 | 5.225 | Glyma20g16920.1 | AT5G61590.1 | ERF/AP2 transcription factor |
| 142 | GmaAffx.15955.1.S1_at | 0.054 | 5.082 | Glyma11g02060.1 | AT2G46410.1 | Similar to CAPRICE, trichoblast fate specification |
| 143 | GmaAffx.64710.1.S1_at | 0.073 | 4.641 | Glyma02g07760.1 | AT5G18270.1 | NAC domain similar to ANAC087 |
| 144 | Gma.4526.1.S1_at | 0.093 | 4.137 | Glyma10g05210.1 | AT2G37430.1 | Zinc finger (C2H2) family protein (ZAT11) |
| 145 | Gma.16807.1.S1_at | 0.040 | 2.685 | Glyma09g14880.1 | AT1G78600.1 | Similar to LZF1, zinc finger (C2H2 type) |
| 146 | GmaAffx.5448.1.S1_at | 0.086 | 2.520 | Glyma16g04720.1 | AT5G18270.1 | NAC domain similar to ANAC087 |
| 147 | GmaAffx.65885.1.A1_s_at | 0.089 | 1.959 | Glyma01g41780.1 | AT2G45120.1 | Zinc finger (C2H2 type) family protein |
| 148 | GmaAffx.9475.1.A1_s_at | 0.039 | 1.845 | Glyma06g14290.1 | AT1G33060.1 | NAC domain similar to ANAC014 |
| 149 | GmaAffx.90028.1.S1_s_at | 0.089 | 1.880 | Glyma04g38560.1 | AT1G01720.1 | NAC domain ATAF1, wound and ABA-induced |
| 150 | Gma.7381.1.S1_at | 0.035 | 1.669 | Glyma20g33430.1 | AT3G10480.1 | NAC domain similar to ANAC050 |
| 151 | Gma.9082.1.S1_at | 0.071 | 1.484 | Glyma15g19980.1 | AT1G19220.1 | Similar to ARF19 |
| 152 | Gma.7454.1.S1_a_at | 0.066 | 2.608 | Glyma13g43800.1 | AT1G04250.1 | Similar to AUX resistant 3 |
| 153 | GmaAffx.39527.1.S1_at | 0.032 | -8.937 | Glyma11g09370.2 | AT2G17410.1 | ARID/BRIGHT DNA-binding domain-containing protein |
| 154 | GmaAffx.4921.1.S1_at | 0.076 | -4.364 | Glyma07g05950.1 | AT3G61850.4 | Similar to DAG1, zinc finger family protein |
| 155 | GmaAffx.70008.1.S1_at | 0.032 | -3.889 | Ambiguous Hit | AT1G66230.1 | Similar to MYB20 |
| 156 | GmaAffx.29616.1.A1_at | 0.069 | -3.371 | Glyma11g04440.2 | AT4G18020.4 | Similar to APRR2 |
| 157 | Gma.8383.1.S1_at | 0.042 | -3.328 | Glyma08g28220.1 | AT3G30530.1 | Similar to ATbZIP42 |
| 158 | Gma.3066.1.S1_at | 0.059 | -3.042 | Glyma19g40070.1 | AT2G47520.1 | Similar to ERF2 subfamily B-2 |
| 159 | GmaAffx.83173.1.S1_at | 0.063 | -2.606 | Glyma14g09310.2 | AT1G46480.1 | Similar to WOX4 |
| 160 | GmaAffx.9569.2.S1_at | 0.074 | -2.207 | Glyma13g20370.2 | AT4G30080.1 | Similar to ARF16, root cap development |
| 161 | GmaAffx.2721.1.S1_s_at | 0.042 | -1.481 | No Soybean Match Identified, E<10E-30, Perecent ID>95% | AT5G37020.1 | Similar to ARF8 |

[0122] Other down-regulated genes included an arid-bright domain protein (GmaAffx.39527.1.S1_at, -8.9 fold). These proteins have been implicated in diverse roles during cell growth and differentiation (Wilsker et al., 2002). A soybean homolog of WUSCHEL-RELATED HOMEOBOX 4 (WOX4) implicated in procambium development was also down-regulated (Ji et al., 2010).

**Example 1 Effects of altered expression of the GmBAG6 gene or GmAP2 in yeast and in plants**

**[0123]** To determine the function of the GmBAG6 and GmAP2 transcription factor identified from the studies described above, Virus Induced Gene Silencing (VIGS) was used to down-regulate the expression of these genes in plants, respectively. VIGS was performed according to Meenu Padmanabhan and Savithramma P. Dinesh-Kumar, "Virus-Induced Gene Silencing as a Tool for Delivery of dsRNA into Plants." Cold Spring Harb. Protoc. (2009), with modifications.

**[0124]** Soybean and Arabidopsis BAG proteins and IQ-BAG domains were expressed in yeast *W303-1A* cells under control of galactose-inducible promoter. The transformed yeast cells were assessed for growth. All yeast cells were cultured in SD-Uracil glucose-based medium to an $OD_{600}$ of about 0.17. Equal numbers of cells were spotted on minimal SD-Uracil medium plates in the presence of glucose (non-inducing) or galactose (inducing). Photographs were taken after culturing at 30°C for 2 days. Fig. 9 shows the results of the these cell death assay, demonstrating that GmBAG6 full length and just its IQ-BAG domains induce cell death when overexpressed in yeast.

**[0125]** Fig. 10 shows range of phenotypes associated with $T_1$ generation of different independent, two week-old *35S* lines overexpressing *GmBag6A* (A-F) and *AtBag6* (G-K) in transgenic Arabidopsis, *Col-0*. (A) Represents wild-type (WT) looking transgenic *35S:GmBag6A* Arabidopsis. (B) Close up image of WT looking transgenic *35S:GmBag6A* Arabidopsis. (C-D) Represents intermediate phenotype associated with *35S:GmBag6A*, sporadic lesions, smaller rosette and leaf malformations. (F) Represents a severe phenotype associated with transgenic *35S:GmBag6A*, stunted plant growth and accumulation of anthocyanin. (G) Represents wild-type (WT) looking transgenic *35S:AtBag6* Arabidopsis. (H) Close up image of WT looking leaf in transgenic *35S:AtBag6* . (I-J) Represents intermediate phenotype associated with *35S:AtBag6*, sporadic lesions, smaller rosette and leaf malformations. (K) Represents a severe phenotype associated with transgenic *35S:AtBag6*, stunted plant growth and accumulation of anthocyanin.

**[0126]** Fig. 11 shows a table summarizing the results of the transgenic Arabidopsis overexpressing the AtBag6 or GmBag6 genes. The percentages of transgenic plants from two separate studies displaying different phenotypes are shown in the table.

**[0127]** BPMV (bean pod mottle virus) was used to overexpress IQ-BAG domain of GmBAG6A (7923R and Glyma07g06750) in soybean plants. For general description of the methodology, see Zhang et al., Plant Physiol. 2010 May; 153(1): 52–65. The plants were

inoculated with either BPMV vector control or IQ-BAG BPMV overexpression construct. As shown in Fig. 12, and by contrast to the vector control, transgenic soybean plants overexpressing the IQ-BAG domain of GmBAG6A display stunting and cell death phenotype. The plants shown are 66-day old.

**[0128]** Fig. 13 shows that VIGS silencing of GmBAG6 (Glyma07g06750) (SEQ ID No. 43) (indicated as (a)) led to increased susceptibility of SCN resistant soybean plants to SCN. This result suggests that GmBAG6 may play a role in SCN resistance. Control plants (EXF67, resistant; EXF63, susceptible) were infected with BPMV with empty vector and are indicated as (b) and (c).

**[0129]** Fig. 14 shows that VIGS silencing of GmAP2 transcription factor (Glyma20g029410) (SEQ ID No. 33) (indicated as (a)) led to increased susceptibility of SCN resistant soybean plants to SCN. This result suggests that GmAP2 may play a role in SCN resistance. Control plants (EXF67, resistant;EXF63, susceptible) were infected with BPMV with empty vector and are indicated as (b) and (c).

**[0130]** In another example, one or more of the SCNRGs, such as GmBAG6 or GmAP2, or their homologs, are placed under a nematode inducible promoter and are introduced into a plant, such as soybean. Expression of the transgene(s) is triggered by the nematode which would confer upon the transgenic plants nematode resistance.

**[0131]** Fig. 15 shows the sequences of GmBAG6A and GmBAG6B as well as their encoded protein sequences. It is to be understood that the materials and methods are taught by way of example, and not by limitation. The disclosed instrumentalities may be broader than the particular methods and materials described herein, which may vary within the skill of the art. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to be limiting.

**[0132]** While the foregoing instrumentalities have been described in some detail for purposes of clarity and understanding, it will be clear to one skilled in the art from a reading of this disclosure that various changes in form and detail can be made without departing from the true scope of the invention. For example, all the techniques and apparatus described above may be used in various combinations. All publications, patents, patent applications, or other documents cited in this application are incorporated by reference in their entirety for all purposes to the same extent as if each individual publication, patent, patent application, or other document were individually indicated to be incorporated by reference for all purposes.

## References

[0133] In addition to those references that are cited in full in the text, additional information for those abbreviated citations is listed below. The content of all patents, patent applications or other publications cited in this disclosure are incorporated by reference into this disclosure.

Acedo JR, Dropkin VH, Luedders VD (1984) Nematode population attrition and histopathology of *Heterodera glycines*-soybean associations. J Nematol 16: 48-56.

Alkharouf NW, Klink VP, Chouikha IB, Beard HS, MacDonald MH, Meyer S, Knap HT, Khan R, Matthews BF (2006) Time course microarray analyses reveal global changes in gene expression of susceptible *Glycine max* (soybean) roots during infection by *Heterodera glycines* (soybean cyst nematode). Planta 224: 838-852.

Balagué C, Lin B, Alcon C, Flottes G, Malmstrom S, Kohler C, Neuhaus G, Pelletier G, Gaymard F, Roby D (2003) HLM1, an essential signaling component in the hypersensitive response, is a member of the cyclic nucleotide-gated channel ion channel family. Plant Cell 15: 365-379.

Bari R, Jones JD (2009) Role of plant hormones in plant defence responses. Plant Mol Biol 69: 473-488.

Branch C, Hwang CF, Navarre DA, Williamson VM (2004) Salicylic acid is part of the *Mi-1*-mediated defense response to root-knot nematode in tomato. Mol Plant Microbe Interact 17: 351-356.

Brucker E, Carlson S, Wright E, Niblack T, Diers B (2005) *Rhg1* alleles from soybean PI 437654 and PI 88788 respond differentially to isolates of *Heterodera glycines* in the greenhouse. Theor Appl Genet 111: 44-49.

Century KS, Holub EB, Staskawicz BJ (1995) NDR1, a locus of *Arabidopsis thaliana* that is required for disease resistance to both a bacterial and a fungal pathogen. Proc Natl Acad Sci U S A 92: 6597-6601.

Century KS, Shapiro AD, Repetti PP, Dahlbeck D, Holub E, Staskawicz BJ (1997) NDR1, a pathogen-induced component required for Arabidopsis disease resistance. Science 278: 1963-1965.

Chiang GC, Barua D, Kramer EM, Amasino RM, Donohue K (2009) Major flowering time gene, flowering locus C, regulates seed germination in *Arabidopsis thaliana*. Proc Natl Acad Sci U S A 106: 11661-11666.

Clough SJ, Fengler KA, Yu IC, Lippok B, Smith RK Jr, Bent AF (2000) The Arabidopsis *dnd1* "defense, no death" gene encodes a mutated cyclic nucleotide-gated ion channel. Proc Natl Acad Sci U S A 97: 9323-9328.

Colgrove AL, Niblack TL (2008) Correlation of female indices from virulence assays on inbred lines and field populations of *Heterodera glycines*. J Nematol 40: 39-45.

Concibido VC , Diers BW, Arelli PR (2004) A decade of QTL mapping for cyst nematode resistance in soybean. Crop Sci 44: 1121-1131.

Coppinger P, Repetti PP, Day B, Dahlbeck D, Mehlert A, Staskawicz BJ (2004) Overexpression of the

plasma membrane-localized NDR1 protein results in enhanced bacterial disease resistance in *Arabidopsis thaliana*. Plant J 40: 225-237.

Cregan PB, Mudge J, Fickus EW, Danesh D, Denny R, Young ND (1999) Two simple sequence repeat markers to select for soybean cyst nematode resistance conditioned by the rhg1 locus. Theorl Appl Genet 99: 811-818.

Duan Y, Zhang W, Li B, Wang Y, Li K, Sodmergen Han C, Zhang Y, Li X (2010) An endoplasmic reticulum response pathway mediates programmed cell death of root tip induced by water stress in Arabidopsis. New Phytol 186: 681-695.

Endo BY (1965) Histological responses of resistant and susceptible soybean varieties, and backcross progeny to entry development of *Heterodera glycines*. Phytopathology 55: 375-381.

Eulgem T, Somssich IE (2007) Networks of WRKY transcription factors in defense signaling. Curr Opin Plant Biol 10: 366-371.

Gao X, Starr J, Gobel C, Engelberth J, Feussner I, Tumlinson J, Kolomiets M (2008) Maize 9-lipoxygenase ZmLOX3 controls development, root-specific expression of defense genes, and resistance to root-knot nematodes. Mol Plant Microbe Interact 21: 98-109.

Glazebrook J (2005) Contrasting mechanisms of defense against biotrophic and necrotrophic pathogens. Annu Rev Phytopathol 43: 205-227.

Grunewald W, Karimi M, Wieczorek K, Van de Cappelle E, Wischnitzki E, Grundler F, Inze D, Beeckman T, Gheysen G (2008) A role for AtWRKY23 in feeding site establishment of plant-parasitic nematodes. Plant Physiol 148: 358-368.

Hubert DA, Tornero P, Belkhadir Y, Krishna P, Takahashi A, Shirasu K, Dangl JL (2003) Cytosolic HSP90 associates with and modulates the Arabidopsis RPM1 disease resistance protein. EMBO J 22: 5679-5689.

Hwang IS, Hwang BK (2010) The pepper 9-lipoxygenase gene *CaLOX1* functions in defense and cell death responses to microbial pathogens. Plant Physiol 152: 948-967.

Ithal N, Recknor J, Nettleton D, Hearne L, Maier T, Baum TJ, Mitchum MG (2007a) Parallel genome-wide expression profiling of host and pathogen during soybean cyst nematode infection of soybean. Mol Plant Microbe Interact 20: 293-305.

Ithal N, Recknor J, Nettleton D, Maier T, Baum TJ, Mitchum MG (2007b) Developmental transcript profiling of cyst nematode feeding cells in soybean roots. Mol Plant Microbe Interact 20: 510-525.

Jia Y, Martin GB (1999) Rapid transcript accumulation of pathogenesis-related genes during an incompatible interaction in bacterial speck disease-resistant tomato plants. Plant Mol Biol 40: 455-465.

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

Journot-Catalino N, Somssich IE, Roby D, Kroj T (2006) The transcription factors WRKY11 and WRKY17 act as negative regulators of basal resistance in *Arabidopsis thaliana*. Plant Cell 18: 3289-3302.

Kang CH, Jung WY, Kang YH, Kim JY, Kim DG, Jeong JC, Baek DW, Jin JB, Lee JY, Kim MO, Chung WS, Mengiste T, Koiwa H, Kwak SS, Bahk JD, Lee SY, Nam JS, Yun DJ, Cho MJ

(2006) AtBAG6, a novel calmodulin-binding protein, induces programmed cell death in yeast and plants. Cell Death Differ 13: 84-95.

Khan R, Alkharouf N, Beard H, Macdonald M, Chouikha I, Meyer S, Grefenstette J, Knap H, Matthews B (2004) Microarray analysis of gene expression in soybean roots susceptible to the soybean cyst nematode two days post invasion. J Nematol 36: 241-248.

Klink VP, Alkharouf N, MacDonald M, Matthews B (2005) Laser capture microdissection (LCM) and expression analyses of *Glycine max* (soybean) syncytium containing root regions formed by the plant pathogen *Heterodera glycines* (soybean cyst nematode). Plant Mol Biol 59: 965-979.

Klink VP, Overall CC, Alkharouf NW, MacDonald MH, Matthews BF (2007a) A time-course comparative microarray analysis of an incompatible and compatible response by *Glycine max* (soybean) to *Heterodera glycines* (soybean cyst nematode) infection. Planta 226: 1423-1447.

Klink VP, Overall CC, Alkharouf NW, MacDonald MH, Matthews BF (2007b) Laser capture microdissection (LCM) and comparative microarray expression analysis of syncytial cells isolated from incompatible and compatible soybean (*Glycine max*) roots infected by the soybean cyst nematode (*Heterodera glycines*). Planta 226: 1389-1409.

Klink VP, Hosseini P, Matsye P, Alkharouf NW, Matthews BF (2009) A gene expression analysis of syncytia laser microdissected from the roots of the *Glycine max* (soybean) genotype PI 548402 (Peking) undergoing a resistant reaction after infection by *Heterodera glycines* (soybean cyst nematode). Plant Mol Biol 71: 525-567.

Klink VP, Hosseini P, Matsye PD, Alkharouf NW, Matthews BF (2010) Syncytium gene expression in *Glycine max* ([PI 88788]) roots undergoing a resistant reaction to the parasitic nematode *Heterodera glycines*. Plant Physiol Biochem 48: 176-193.

Koenning SR, Wrather JA (2010). Suppression of soybean yield potential in the continental United States from plant diseases estimated from 2006 to 2009. Plant Health Progress, in press.

Lamb C, Dixon RA (1997) The oxidative burst in plant disease resistance. Annu Rev Plant Physiol Plant Mol Biol 48: 251-275.

Lee MW, Lu H, Jung HW, Greenberg JT (2007) A key role for the Arabidopsis WIN3 protein in disease resistance triggered by *Pseudomonas syringae* that secrete AvrRpt2. Mol Plant Microbe Interact 20: 1192-1200.

Lee MW, Jelenska J, Greenberg JT (2008) Arabidopsis proteins important for modulating defense responses *to Pseudomonas syringae* that secrete HopW1-1. Plant J 54: 452-465.

Li Y, Chen S, Young ND (2004) Effect of *rhg1* gene on penetration, development, and reproduction of *Heterodera glycines* race 3. Nematology 6: 727-734.

Mahalingam R, Skorupska HT (1996) Cytological expression of early response to infection by *Heterodera glycines* Ichinohe in resistant PI 437654 soybean. Genome 39: 986-998.

Manickavelu A, Kawaura K, Oishi K, Shin IT, Kohara Y, Yahiaoui N, Keller B, Suzuki A, Yano K, Ogihara Y (2010) Comparative gene expression analysis of susceptible and resistant near-isogenic lines in common wheat infected by *Puccinia triticina*. DNA Res 17: 211-222.

Meier S, Bastian R, Donaldson L, Murray S, Bajic V, Gehring C (2008) Co-expression and promoter content analyses assign a role in biotic and abiotic stress responses to plant natriuretic peptides. BMC Plant Biol 8: 24.

Melito S, Heuberger AL, Cook D, Diers BW, MacGuidwin AE, Bent AF (2010) A nematode demographics assay in transgenic roots reveals no significant impacts of the *Rhg1* locus LRR-Kinase on soybean cyst nematode resistance. BMC Plant Biol 10: 104.

Morel JB, Dangl JL (1997) The hypersensitive response and the induction of cell death in plants. Cell Death Differ 4: 671-683.

Morse M, Pironcheva G, Gehring C (2004) AtPNP-A is a systemically mobile natriuretic peptide immunoanalogue with a role in *Arabidopsis thaliana* cell volume regulation. FEBS Lett 556: 99-103.

Mudge J (1999) Marker assisted selection for soybean cyst nematode resistance and accompanying agronomic traits Ph.D. thesis University of Minnesota St. Paul, MN

Niblack TL, Arelli PR, Noel GR, Opperman CH, Orf JH, Schmitt DP, Shannon JG, Tylka GL (2002) A revised classification scheme for genetically diverse populations of *Heterodera glycines*. J. Nematol 34: 279-288.

Niblack TL, Heinz RD, Smith GS, Donald PA (1993) Distribution, density, and diversity of Heterodera glycines in Missouri. J. Nematol 25(4S): 880-886.

Nishizawa A, Yabuta Y, Yoshida E, Maruta T, Yoshimura K, Shigeoka S (2006) Arabidopsis heat shock transcription factor A2 as a key regulator in response to several types of environmental stress. Plant J 48: 535-547.

Nobuta K, Okrent RA, Stoutemyer M, Rodibaugh N, Kempema L, Wildermuth MC, Innes RW (2007) The GH3 acyl adenylase family member PBS3 regulates salicylic acid-dependent defense responses in Arabidopsis. Plant Physiol 144: 1144-1156.

O'Rourke JA, Nelson RT, Grant D, Schmutz J, Grimwood J, Cannon S, Vance CP, Graham MA, Shoemaker RC (2009) Integrating microarray analysis and the soybean genome to understand the soybeans iron deficiency response. BMC Genomics 10: 376.

Onishi M, Tachi H, Kojima T, Shiraiwa M, Takahara H (2006) Molecular cloning and characterization of a novel salt-inducible gene encoding an acidic isoform of PR-5 protein in soybean (*Glycine max [L.] Merr.*). Plant Physiol Biochem 44: 574-580.

Passardi F, Penel C, Dunand C (2004) Performing the paradoxical: how plant peroxidases modify the cell wall. Trends Plant Sci 9: 534-540.

Potter S, Uknes S, Lawton K, Winter A-M, Chandler D, DiMaio J, Novitzky R, Ward E, Ryals J (1993) Regulation of a hevein-like gene in Arabidopsis. Mol Plant-Microbe Interact 6: 680-685.

Raffaele S, Rivas S, Roby D (2006) An essential role for salicylic acid in AtMYB30-mediated control of the hypersensitive cell death program in Arabidopsis. FEBS Lett 580: 3498-3504.

Riggs RD, Kim KS, Gipson I (1973) Ultrastructural changes in Peking soybeans infected with *Heterodera glycines*. Phytopathology 63: 76-84.

Ross JP (1958) Host-parasite relationship of the soybean cyst nematode in resistant soybean roots. Phytopathology 48: 578-579.

Schlueter JA, Dixon P, Granger C, Grant D, Clark L, Doyle JJ, Shoemaker RC (2004) Mining EST databases to resolve evolutionary events in major crop species. Genome 47: 868-876.

Schlueter JA, Lin JY, Schlueter SD, Vasylenko-Sanders IF, Deshpande S, Yi J, O'Bleness M, Roe BA, Nelson RT, Scheffler BE, Jackson SA, Shoemaker RC (2007) Gene duplication and paleopolyploidy in soybean and the implications for whole genome sequencing. BMC Genomics 8: 330.

Schmutz J, Cannon SB, Schlueter J, Ma J, Mitros T, Nelson W, Hyten DL, Song Q, Thelen JJ, Cheng J, Xu D, Hellsten U, May GD, Yu Y, Sakurai T, Umezawa T, Bhattacharyya MK, Sandhu D, Valliyodan B, Lindquist E, Peto M, Grant D, Shu S, Goodstein D, Barry K, Futrell-Griggs M, Abernathy B, Du J, Tian Z, Zhu L, Gill N, Joshi T, Libault M, Sethuraman A, Zhang XC, Shinozaki K, Nguyen HT, Wing RA, Cregan P, Specht J, Grimwood J, Rokhsar D, Stacey G, Shoemaker RC, Jackson SA (2010) Genome sequence of the palaeopolyploid soybean. Nature 463: 178-183.

Shapiro AD, Zhang C (2001) The role of NDR1 in avirulence gene-directed signaling and control of programmed cell death in Arabidopsis. Plant Physiol 127:1089-1101.

Storey JD, Tibshirani R (2003) Statistical significance for genomewide studies. Proc Natl Acad Sci U S A 100: 9440-9445.

Takahashi A, Casais C, Ichimura K, Shirasu K (2003) HSP90 interacts with RAR1 and SGT1 and is essential for RPS2-mediated disease resistance in Arabidopsis. Proc Natl Acad Sci U S A 100: 11777-11782.

Urade R (2009) The endoplasmic reticulum stress signaling pathways in plants. Biofactors 35: 326-331.

Vailleau F, Daniel X, Tronchet M, Montillet JL, Triantaphylides C, Roby D (2002) A R2R3-MYB gene, AtMYB30, acts as a positive regulator of the hypersensitive cell death program in plants in response to pathogen attack. Proc Natl Acad Sci U S A 99: 10179-10184.

Vanderbeld B, Snedden WA (2007) Developmental and stimulus-induced expression patterns of Arabidopsis calmodulin-like genes CML37, CML38 and CML39. Plant Mol Biol 64: 683-697.

Varet A, Parker J, Tornero P, Nass N, Numberger T, Dangl JL, Scheel D, Lee J (2002) NHL25 and NHL3, two NDR1/HIN1-like genes in *Arabidopsis thaliana* with potential role(s) in plant defense. Mol Plant Microbe Interact 15: 608-616.

Varet A, Hause B, Hause G, Scheel D, Lee J (2003) The Arabidopsis *NHL3* gene encodes a plasma membrane protein and its overexpression correlates with increased resistance to *Pseudomonas syringae pv.* tomato DC3000. Plant Physiol 132: 2023-2033.

Wang W, Vinocur B, Shoseyov O, Altman A (2004) Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response. Trends Plant Sci 9: 244-252.

Wang X, Replogle A, Davis EL, Mitchum MG (2007) The tobacco *Cel7* gene promoter is auxin-responsive and locally induced in nematode feeding sites of heterologous plants. Mol Plant Pathol 8: 423-436.

Xiao YL, Smith SR, Ishmael N, Redman JC, Kumar N, Monaghan EL, Ayele M, Haas BJ, Wu HC, Town CD (2005) Analysis of the cDNAs of hypothetical genes on Arabidopsis chromosome 2 reveals numerous transcript variants. Plant Physiol 139: 1323-1337.

Zhang C, Yang C, Whitham SA, Hill JH (2009) Development and use of an efficient DNA-based viral gene silencing vector for soybean. Mol Plant Microbe Interact 22: 123-131.

Zhang C, Bradshaw JD, Whitham SA, Hill JH (2010) The development of an efficient multipurpose

bean pod mottle virus viral vector set for foreign gene expression and RNA silencing. Plant Physiol 153: 52-65.

Zheng MS, Takahashi H, Miyazaki A, Hamamoto H, Shah J, Yamaguchi I, Kusano T (2004) Up-regulation of *Arabidopsis thaliana* NHL10 in the hypersensitive response to Cucumber mosaic virus infection and in senescing leaves is controlled by signalling pathways that differ in salicylate involvement. Planta 218: 740-750.

Zheng Z, Qamar SA, Chen Z, Mengiste T (2006) Arabidopsis WRKY33 transcription factor is required for resistance to necrotrophic fungal pathogens. Plant J 48: 592-605.

Chao, Q., Rothenberg, M., Solano, R., Roman, G., Terzaghi, W., Ecker, J.R. (1997) Activation of the ethylene gas response pathway in arabidopsis by the nuclear protein ETHYLENE-INSENSITIVE3 and related proteins Cell, 89 (7), pp. 1133-1144

Gechev, T.S., Minkov, I.N., Hille, J.(2005) Hydrogen peroxide-induced cell death in arabidopsis: Transcriptional and mutant analysis reveals a role of an oxoglutarate-dependent dioxygenase gene in the cell death process IUBMB Life 57 (3), pp. 181-188

Ji, J., Strable, J., Shimizu, R., Koenig, D., Sinha, N., Scanlon, M.J. (2010) Wox4 promotes procambial development Plant Physiology 152 (3), pp. 1346-1356

Nagpal, P., Ellis, C.M., Weber, H., Ploense, S.E., Barkawi, L.S., Guilfoyle, T.J., Hagen, G., (...), Reed, J.W. (2005) Auxin response factors ARF6 and ARF8 promote jasmonic acid production and flower maturation Development 132 (18), pp. 4107-4118

Wang, X., Basnayake, B.M.V.S., Zhang, H., Li, G., Li, W., Virk, N., Mengiste, T., Song, F.(2009) The Arabidopsis ATAF1, a NAC transcription factor, is a negative regulator of defense responses against neurotrophic fungal and bacterial pathogens Molecular Plant-Microbe Interactions 22 (10), pp. 1227-1238

Wilsker, D., Patsialou, A., Dallas, P.B., Moran, E.(2002) ARID proteins: A diverse family of DNA binding proteins implicated in the control of cell growth, differentiation, and development Cell Growth and Differentiation 13 (3), pp. 95-106

Zhong, R., Lee, C., Zhou, J., McCarthy, R.L., Ye, Z.-H. (2008) A battery of transcription factors involved in the regulation of secondary cell wall biosynthesis in Arabidopsis Plant Cell 20 (10), pp. 2763-2782.

## Claims

We claim:

1. A transgenic plant generated from a host plant, said transgenic plant being generated by altering the expression levels of a protein encoded by a Soybean Cyst Nematode Resistant Gene (SCNRG) or a fragment thereof, wherein said transgenic plant is more resistant to nematode infection when compared to the host plant.

2. The transgenic plant of claim 1, wherein said SCNRG is endogenous to the host plant.

3. The transgenic plant of claim 1, wherein said SCNRG is exogenous to the host plant and said SCNRG is introduced into said host plant.

4. The transgenic plant of any one of the preceding claims, wherein said host plant is a soybean plant, and said host plant is susceptible to soybean cyst nematode (SCN) infection.

5. The transgenic plant of any one of the preceding claims, wherein the expression level of said protein encoded by said SCNRG is higher in said transgenic plant than the expression level of said protein in the host plant.

6. The transgenic plant of any one of the preceding claims, wherein the expression level of said protein is at least two fold higher in said transgenic plant than the expression level of said protein in the host plant.

7. The transgenic plant of any one of the preceding claims, wherein the expression of said SCNRG is under control of a nematode inducible promoter.

8. The transgenic plant of any one of the preceding claims, wherein the expression level of said protein encoded by said SCNRG is lower in said transgenic plant than the expression level of said protein in the host plant.

9. The transgenic plant of any one of the preceding claims, wherein the expression levels of two more proteins encoded by two or more SCNRGs are altered in the transgenic plant.

10. The transgenic plant of any one of the preceding claims, wherein the SCNRG is selected from the group consisting of GmBAG6, GmAP2, GmBAG6 homolog, GmAP2 homolog, and combination thereof.

11. A method for generating a transgenic plant using a host plant, said transgenic plant being more resistant to soybean cyst nematode (SCN) infection when compared to the host plant, said method comprising a step of altering the expression levels of a protein encoded by a Soybean Cyst Nematode Resistant Gene (SCNRG) or a fragment thereof, said SCNRG being endogenous to the host plant.

12. The method of any one of the preceding claims, wherein the expression level of said protein encoded by said SCNRG is higher in said transgenic plant than the expression level of said SCNRG in the host plant.

13. The method of any one of the preceding claims, wherein the expression level of said protein encoded by said SCNRG is lower in said transgenic plant than the expression level of said SCNRG in the host plant.

14. The method of any one of the preceding claims, wherein the level of said protein encoded by said SCNRG is altered by gene silencing.

15. A method for generating a transgenic plant using a host plant, said transgenic plant being more resistant to soybean cyst nematode (SCN) infection when compared to the host plant, said method comprising a step of introducing Soybean Cyst Nematode Resistant Gene (SCNRG) or a fragment thereof into said host plant, wherein the SCNRG causes the expression of a protein encoded by said SCNRG, said protein being capable of rendering said host plant more resistant to SCN infection.

16. The method of any one of the preceding claims, wherein said host plant is a soybean plant, and said host plant is susceptible to soybean cyst nematode (SCN) infection.

17. The method of claim any one of the preceding claims, wherein the expression of said SCNRG is under control of a nematode inducible promoter.

18. The method of any one of the preceding claims, wherein the expression levels of two more proteins encoded by two or more SCNRGs are altered in the transgenic plant.

19. The method of any one of the preceding claims, wherein the SCNRG is selected from the group consisting of GmBAG6, GmAP2, GmBAG6 homolog, GmAP2 homolog, and combination thereof.

20. The method of any one of the preceding claims, wherein the SCNRG is selected from the group consisting of SEQ ID Nos. 1-169, and their homologs.

21. The transgenic plant of any one of the preceding claims, wherein the SCNRG is selected from the group consisting of SEQ ID Nos. 1-169, and their homologs.

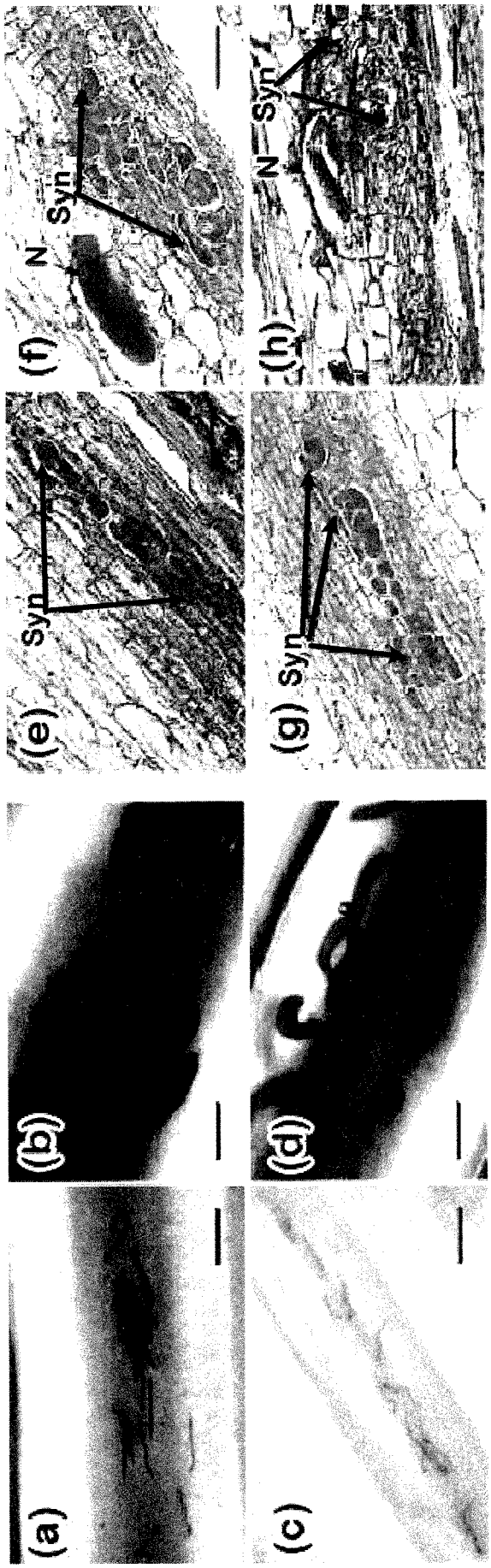



FIG. 1

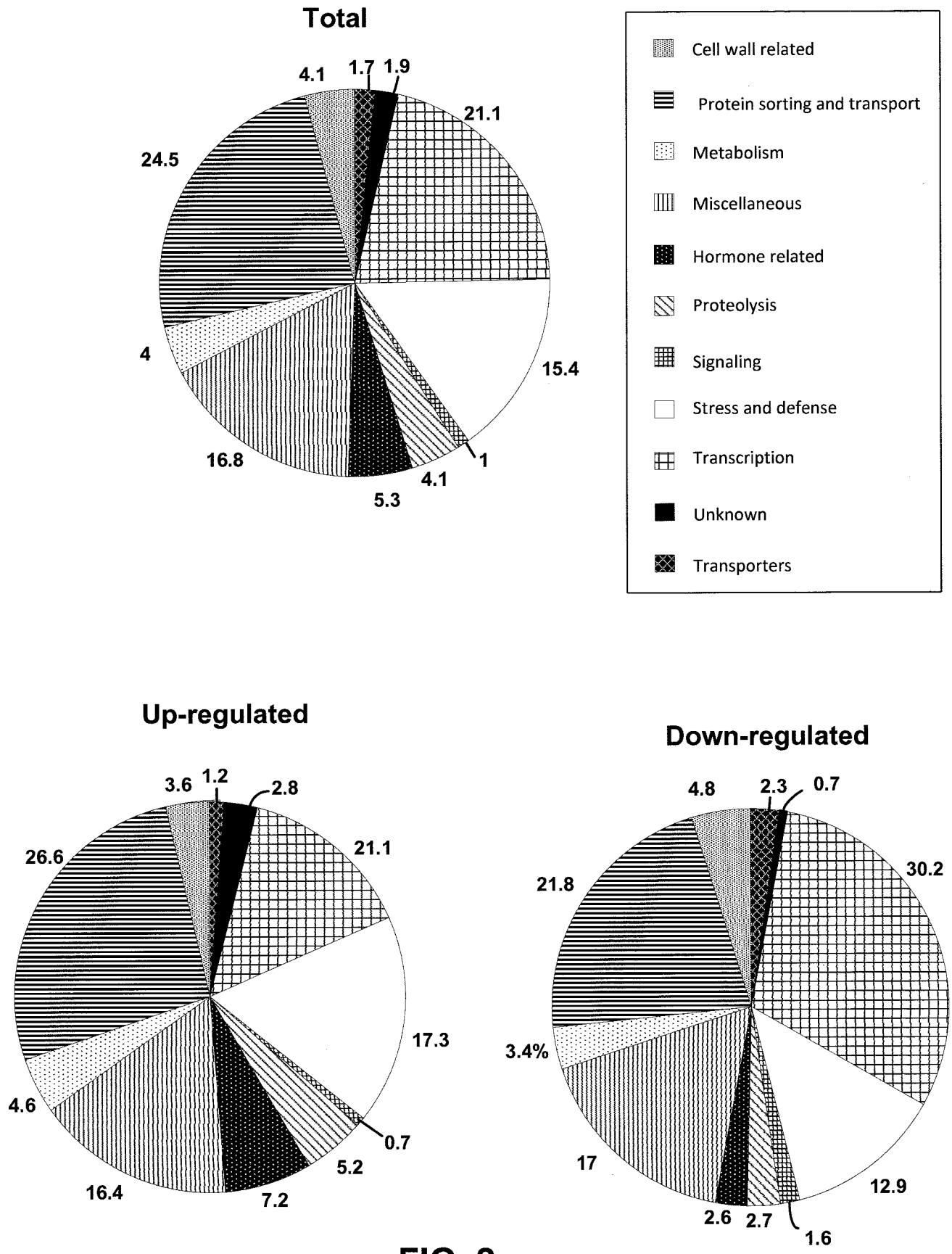

Total
4.1
1.7
1.9
21.1
24.5
4
15.4
16.8
1
4.1
5.3
Cell wall related
Protein sorting and transport
Metabolism
Miscellaneous
Hormone related
Proteolysis
Signaling
Stress and defense
Transcription
Unknown
Transporters
Up-regulated
3.6
1.2
2.8
21.1
26.6
17.3
4.6
0.7
5.2
16.4
7.2
Down-regulated
4.8
2.3
0.7
30.2
21.8
3.4%
17
12.9
2.6
2.7
1.6

FIG. 2

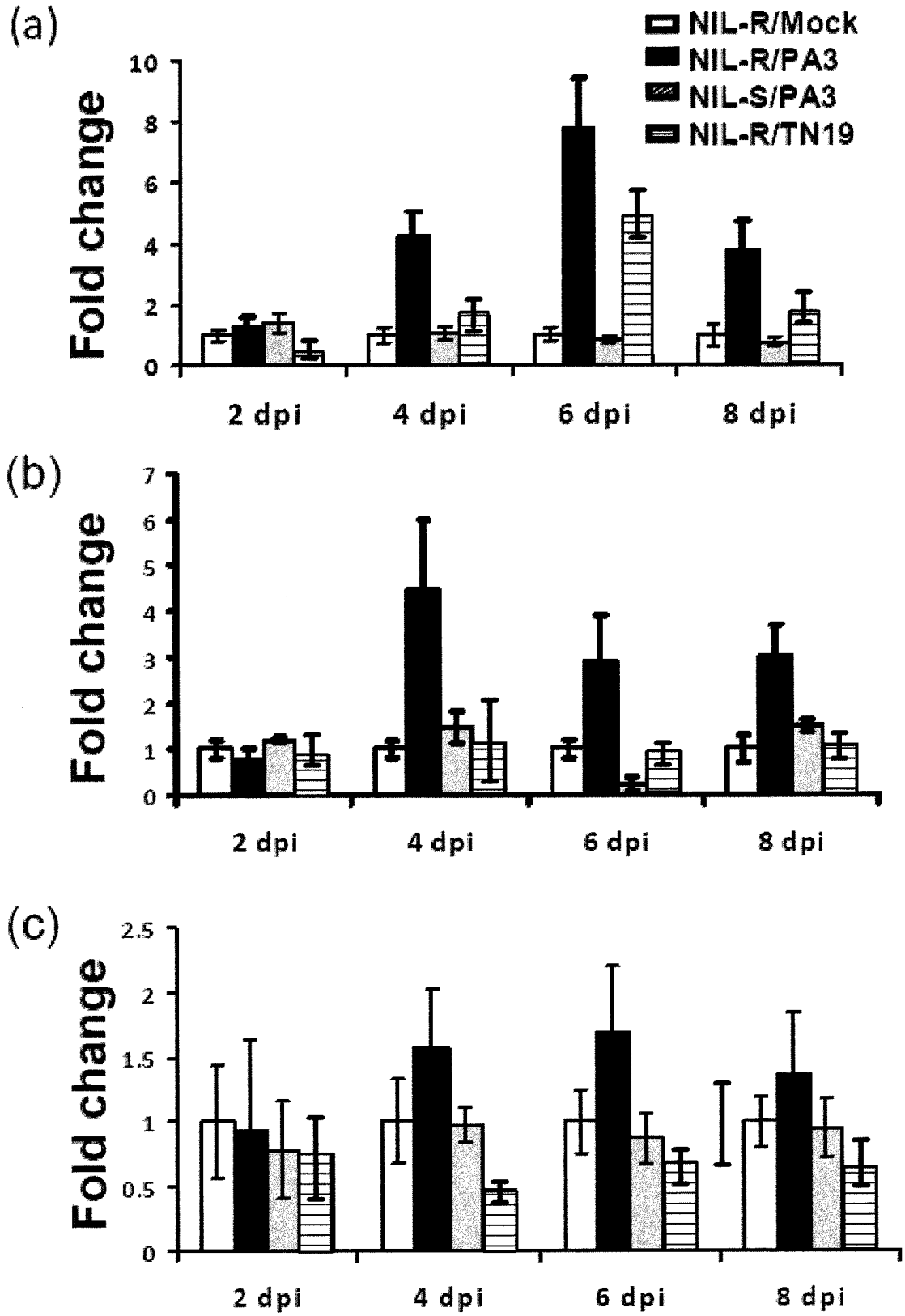

(a)
NIL-R/Mock
NIL-R/PA3
NIL-S/PA3
NIL-R/TN19
Fold change
10
8
6
4
2
0
2 dpi
4 dpi
6 dpi
8 dpi
(b)
Fold change
7
6
5
4
3
2
1
0
2 dpi
4 dpi
6 dpi
8 dpi
(c)
Fold change
2.5
2
1.5
1
0.5
0
2 dpi
4 dpi
6 dpi
8 dpi

FIG. 3

| Affychip ID | Gene Model # | Fold-Upregulation (Microarray) | Description |
|---|---|---|---|
| ArWRKY23 | At2g47260 | N/A | WRKY transcription factor, induced in syncytia by *H. Schachtii* in Arabidopsis |
| Gma.986.1.S1_at | Glyma15g04570.1 | 11.1 | Zinc finger (C2H2) transcription factor |
| Gma.15997.1.S1_at | Glyma15g06130.1 | 9.9 | Arabinogalactan protein |
| Gma.9553.1.A1_at | Glyma14g06080.1 | 8.08 | AP2 transcription factor, DREB family A2 |
| Gma.487.1.S1_at | Glyma18g43750.1 | 6.86 | AP2 transcription factor, DREB2B |
| GmaAffx.2469.1.S1_at | Glyma01g42500.2 | 6.92 | AP2 transcription factor, DDF1 |
| Gma.10412.1.A1_at | Glyma01g42440.1 | 5.92 | Nucleotide rhamonose synthase NRS/ER |
| GmaAffx.80271.1.S1_at | Glyma13g35100.1 | 3.77 | Extensin-like protein involved in seed germination |
| Gma.7147.1.A1_at | Glyma09g34110.1 | 2.98 | Acyltransferase 1, ATS1 |
| GmaAffx.7387.1.S1_at | Glyma18g18060.1 | 2.36 | Xanthine/uracil permease family protein |
| GmaAffx.45036.1.S1_at | Glyma19g04410.1 | 2.35 | Pollen protein, Ole e 1 family protein |

FIG. 4A

(b)

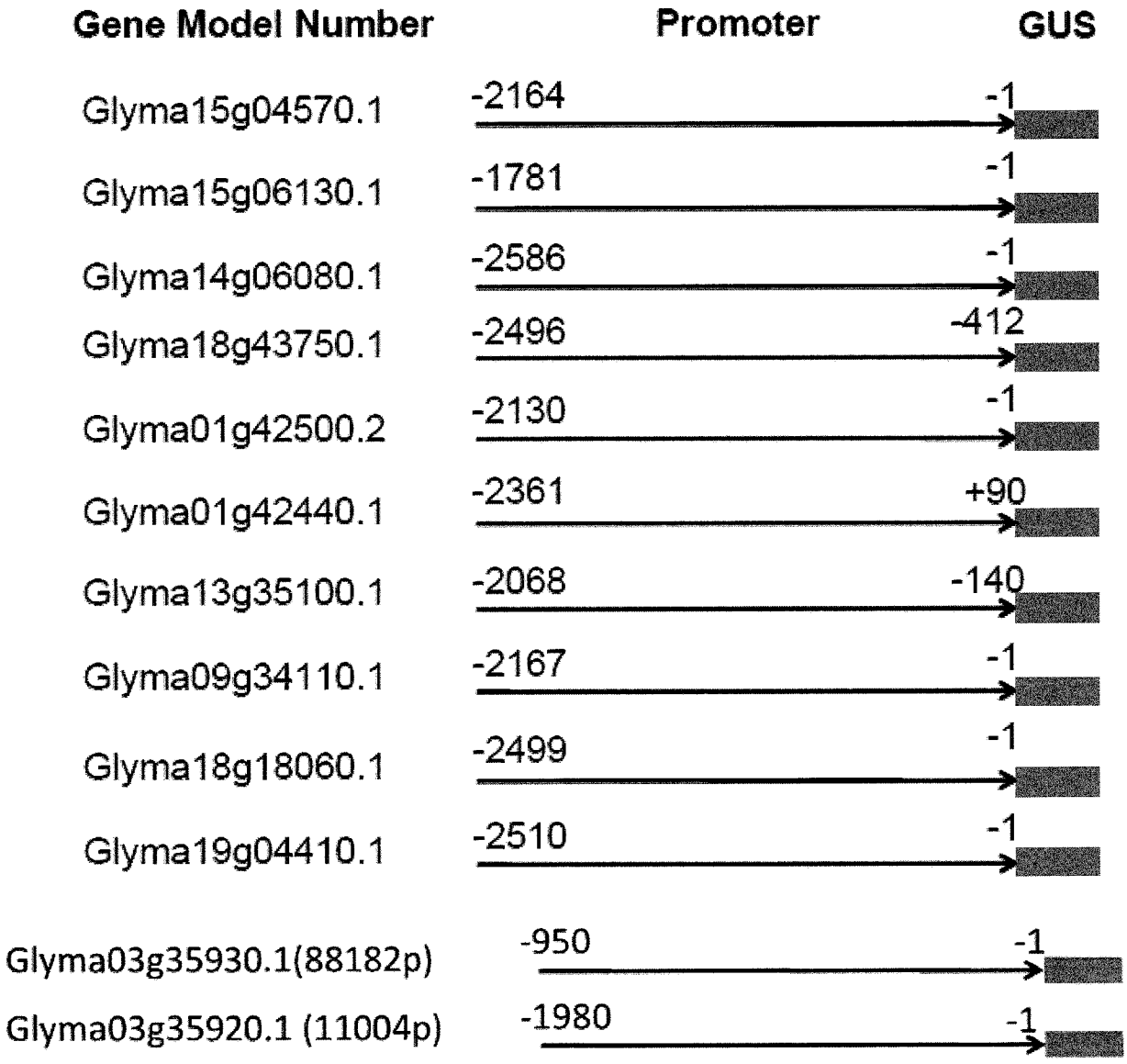

Gene Model Number
Promoter
GUS
Glyma15g04570.1
-2164
-1
Glyma15g06130.1
-1781
-1
Glyma14g06080.1
-2586
-1
Glyma18g43750.1
-2496
-412
Glyma01g42500.2
-2130
-1
Glyma01g42440.1
-2361
+90
Glyma13g35100.1
-2068
-140
Glyma09g34110.1
-2167
-1
Glyma18g18060.1
-2499
-1
Glyma19g04410.1
-2510
-1
Glyma03g35930.1(88182p)
-950
-1
Glyma03g35920.1 (11004p)
-1980
-1

FIG. 4B

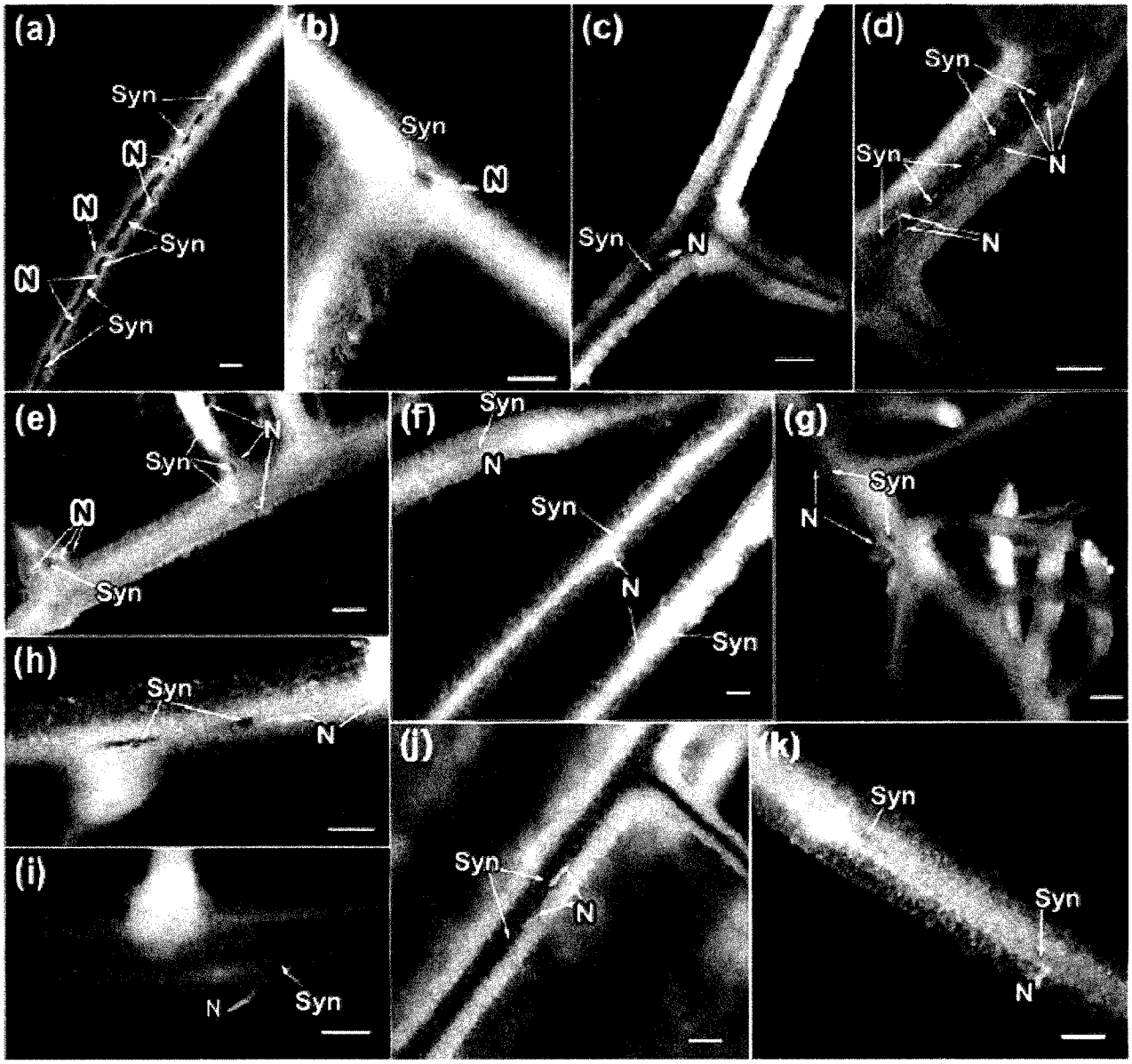

(a)
Syn
N
N
Syn
N
Syn
(b)
Syn
N
(c)
Syn
N
(d)
Syn
Syn
N
N
(e)
N
Syn
N
Syn
(f)
Syn
N
Syn
N
Syn
(g)
Syn
N
(h)
Syn
N
(i)
N
Syn
(j)
Syn
N
(k)
Syn
Syn
N

FIG. 5

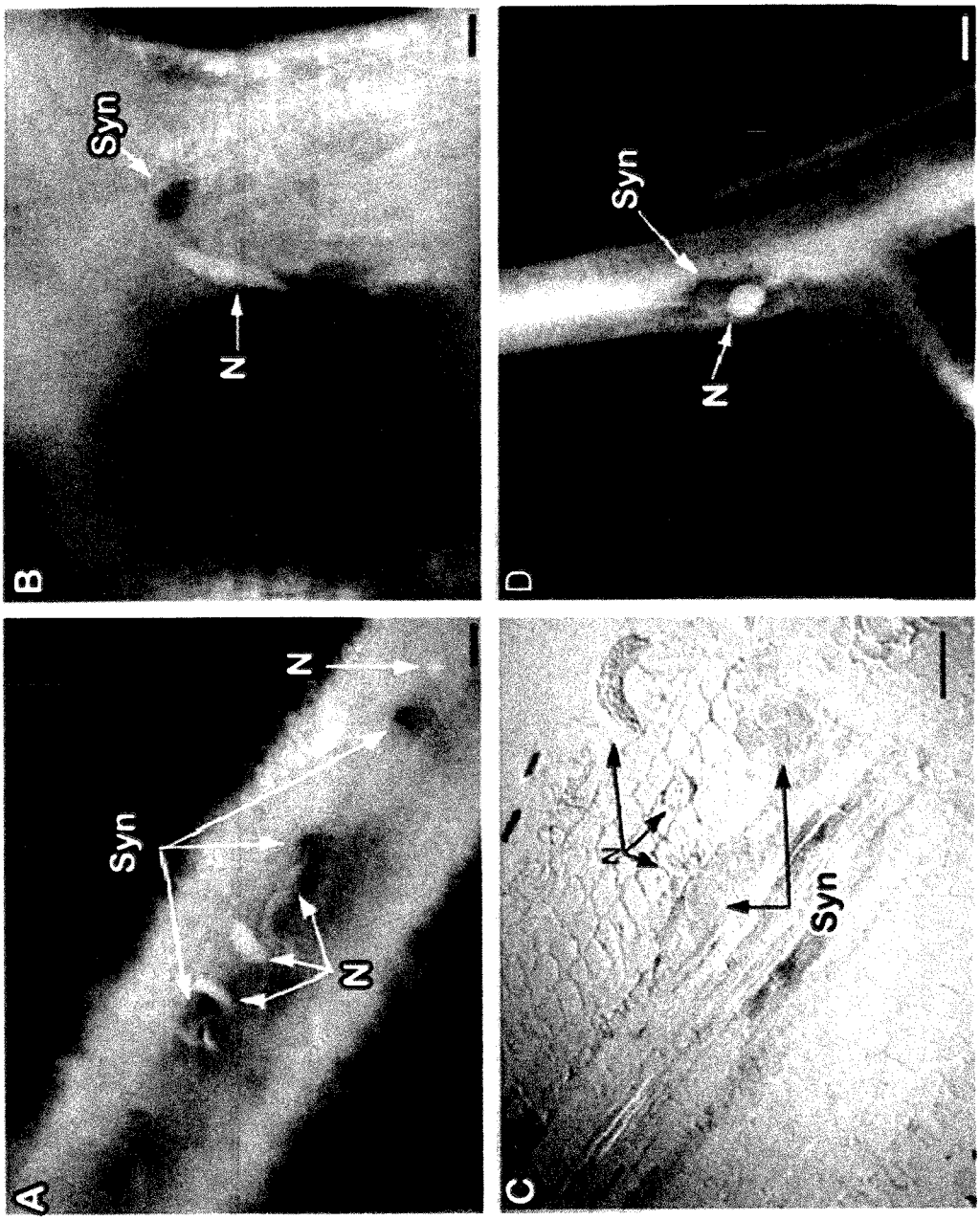



FIG. 6

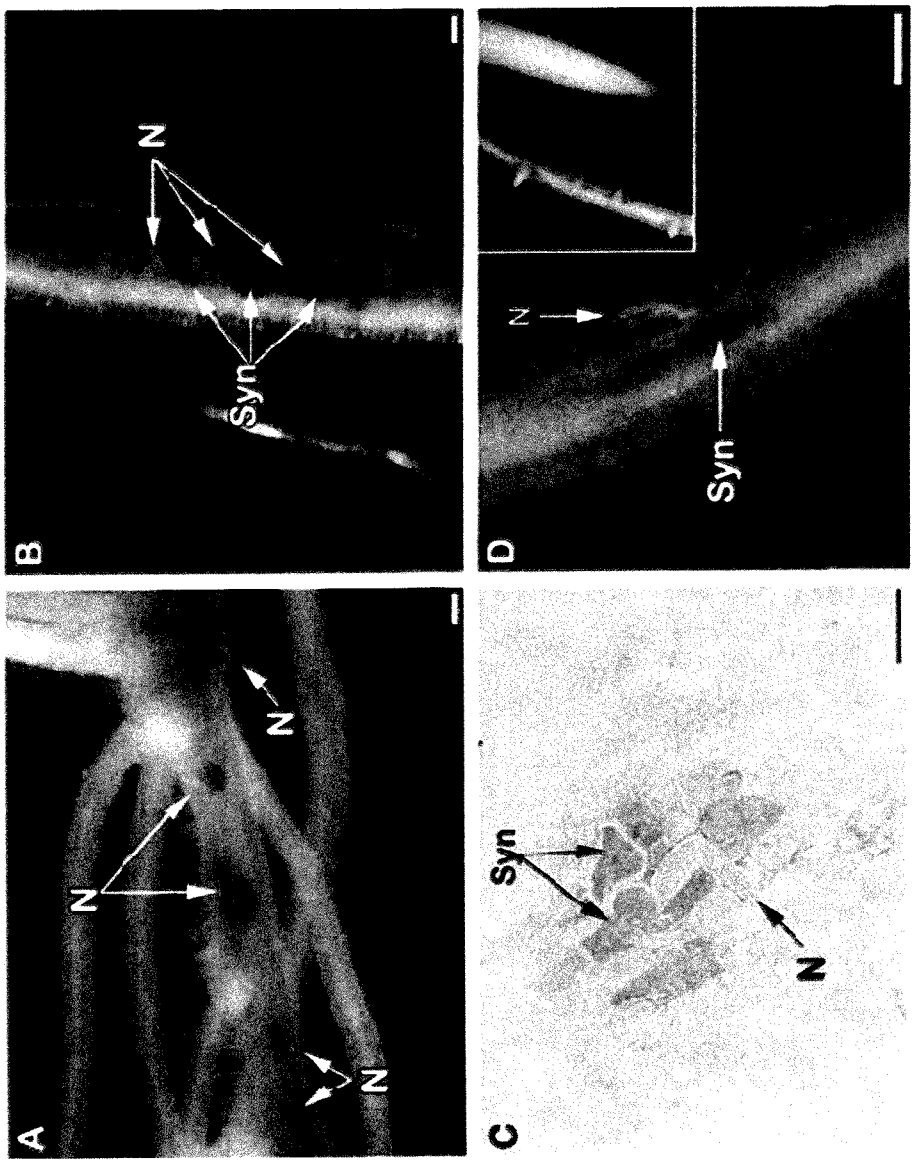



FIG. 7

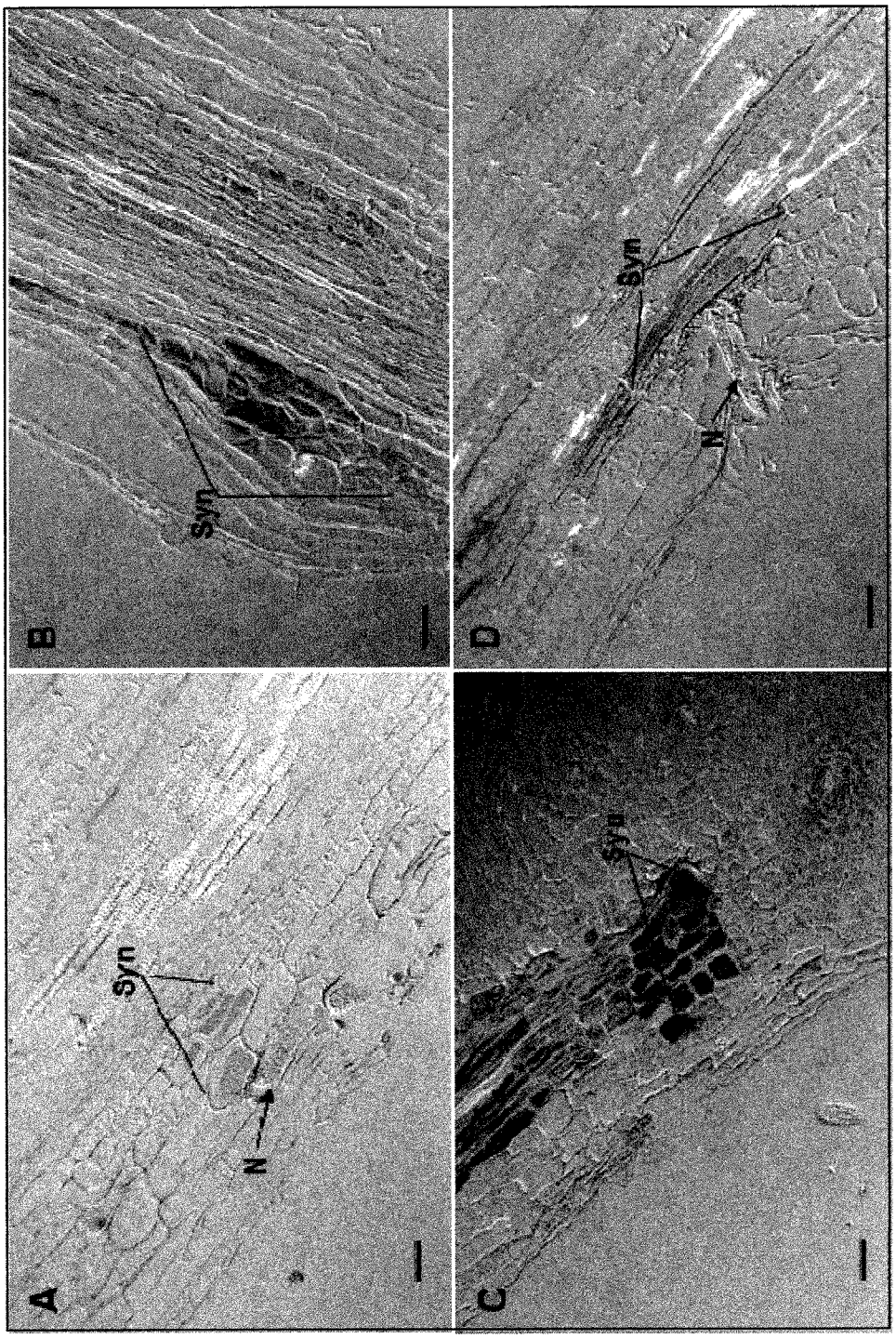



FIG. 8

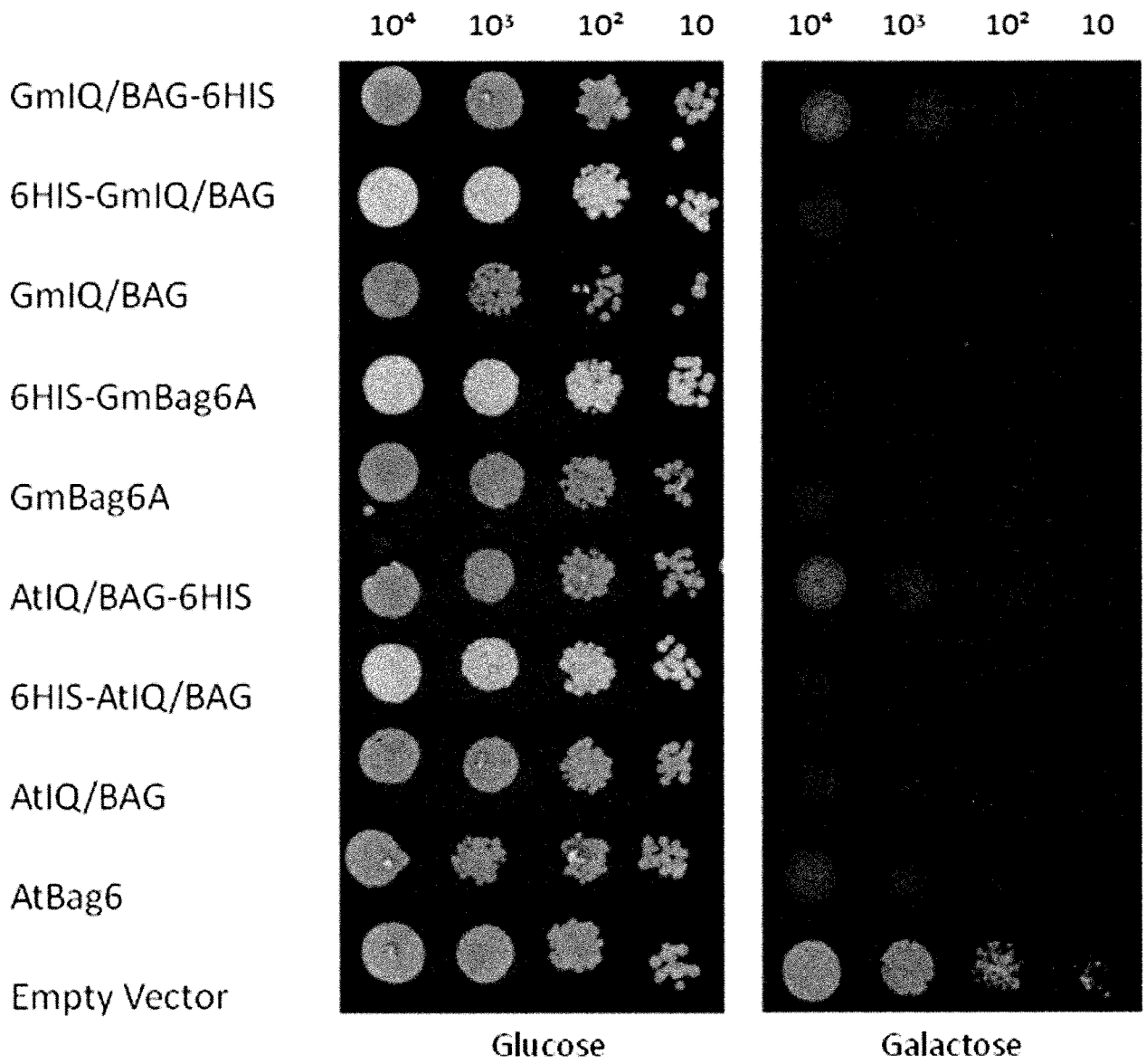

10⁴ 10³ 10² 10
10⁴ 10³ 10² 10
GmIQ/BAG-6HIS
6HIS-GmIQ/BAG
GmIQ/BAG
6HIS-GmBag6A
GmBag6A
AtIQ/BAG-6HIS
6HIS-AtIQ/BAG
AtIQ/BAG
AtBag6
Empty Vector
Glucose
Galactose

FIG. 9

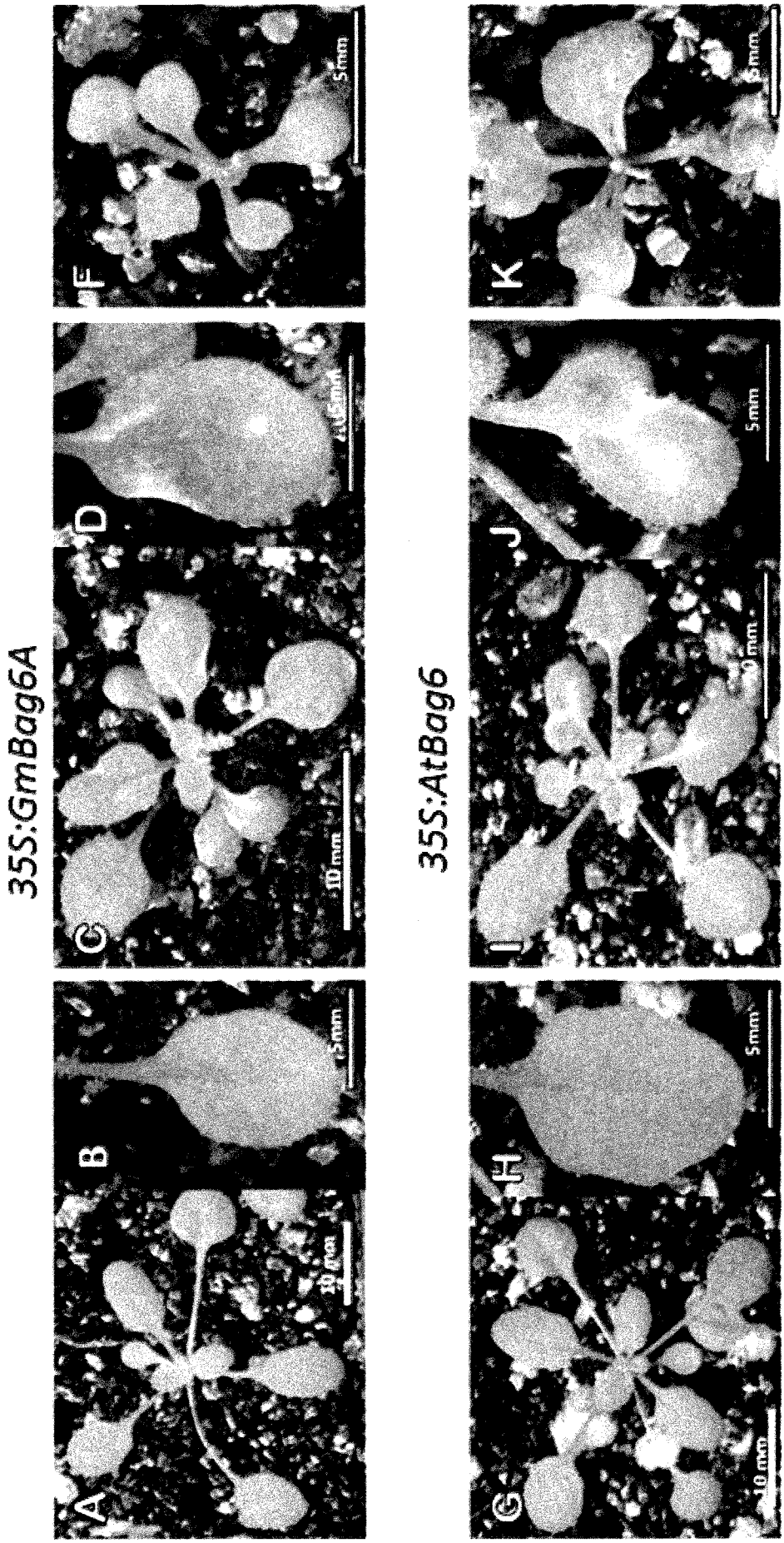



FIG. 10

Table 1. Summary of *BAG6* Overexpression Phenotypes in Arabidopsis

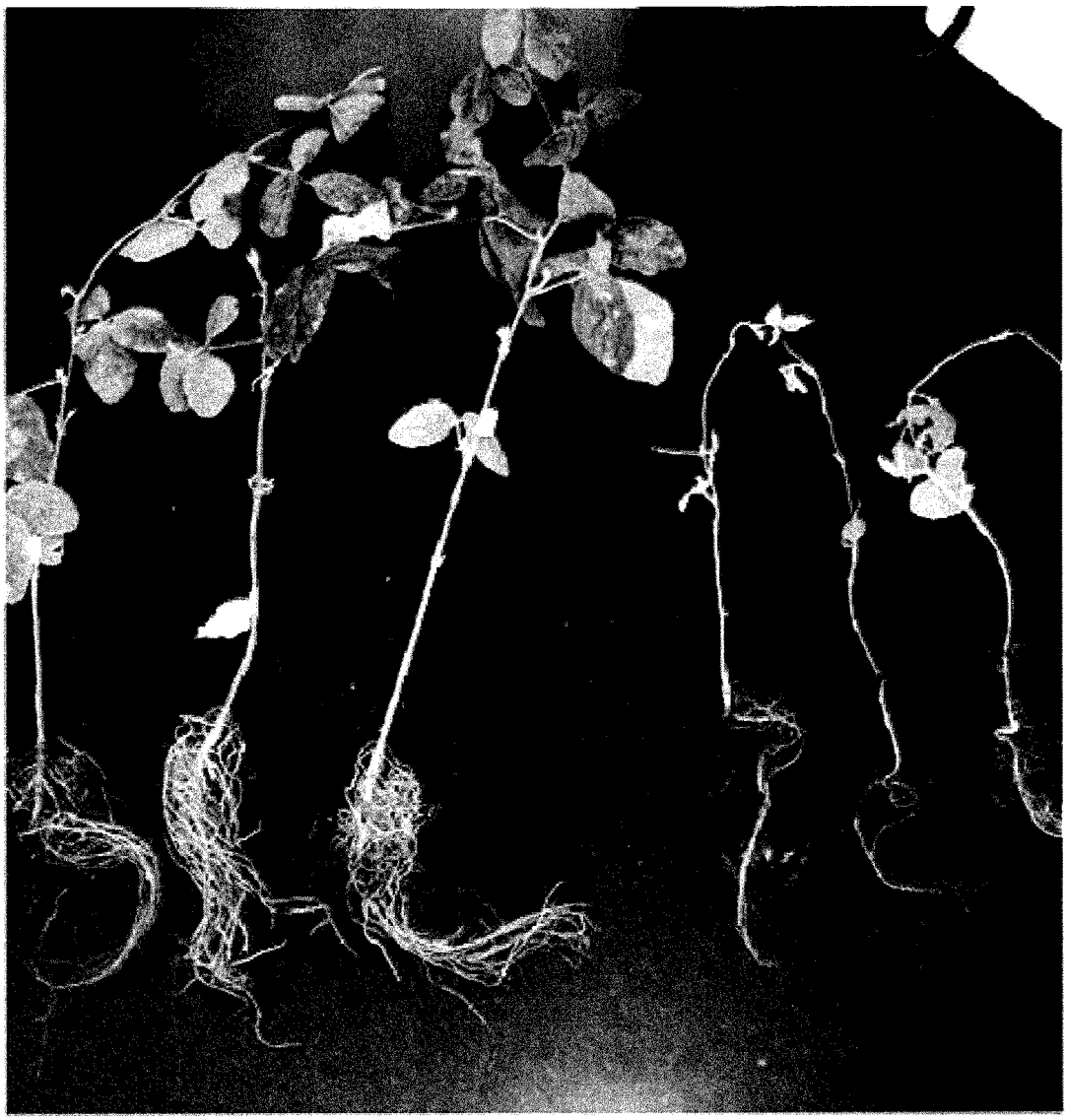
| | $T_1$ Shoot Phenotypes[a] | | | | |
|---|---|---|---|---|---|
| Construct | WT-like (%) | Intermediate (%) | Severe (%) | $T_1$ lines (no.) | Total $T_1$ (no.) |
| *35S:AtBag6* | 43.9 | 40.9 | 15.2 | 6 | 66 |
| *35S:GmBag6A* | 42 | 52 | 6 | 12 | 100 |

| | $T_1$ Shoot Phenotypes[a] | | | | |
|---|---|---|---|---|---|
| Construct | WT-like (%) | Intermediate (%) | Severe (%) | $T_1$ lines (no.) | Total $T_1$ (no.) |
| *35S:AtBag6* | 42 | 46 | 12 | 8 | 50 |
| *35S:GmBag6A* | 38 | 50 | 12 | 8 | 66 |

[a]2-week-old seedlings; WT-like = indistinguishable from wildtype; intermediate = smaller rosettes, leaf lesions, malformed leaves; severe = dwarf, anthocyanin accumulation



FIG. 11



BPMV Control Plants

BPMV IQ-BAG domain overexpression plants

FIG. 12

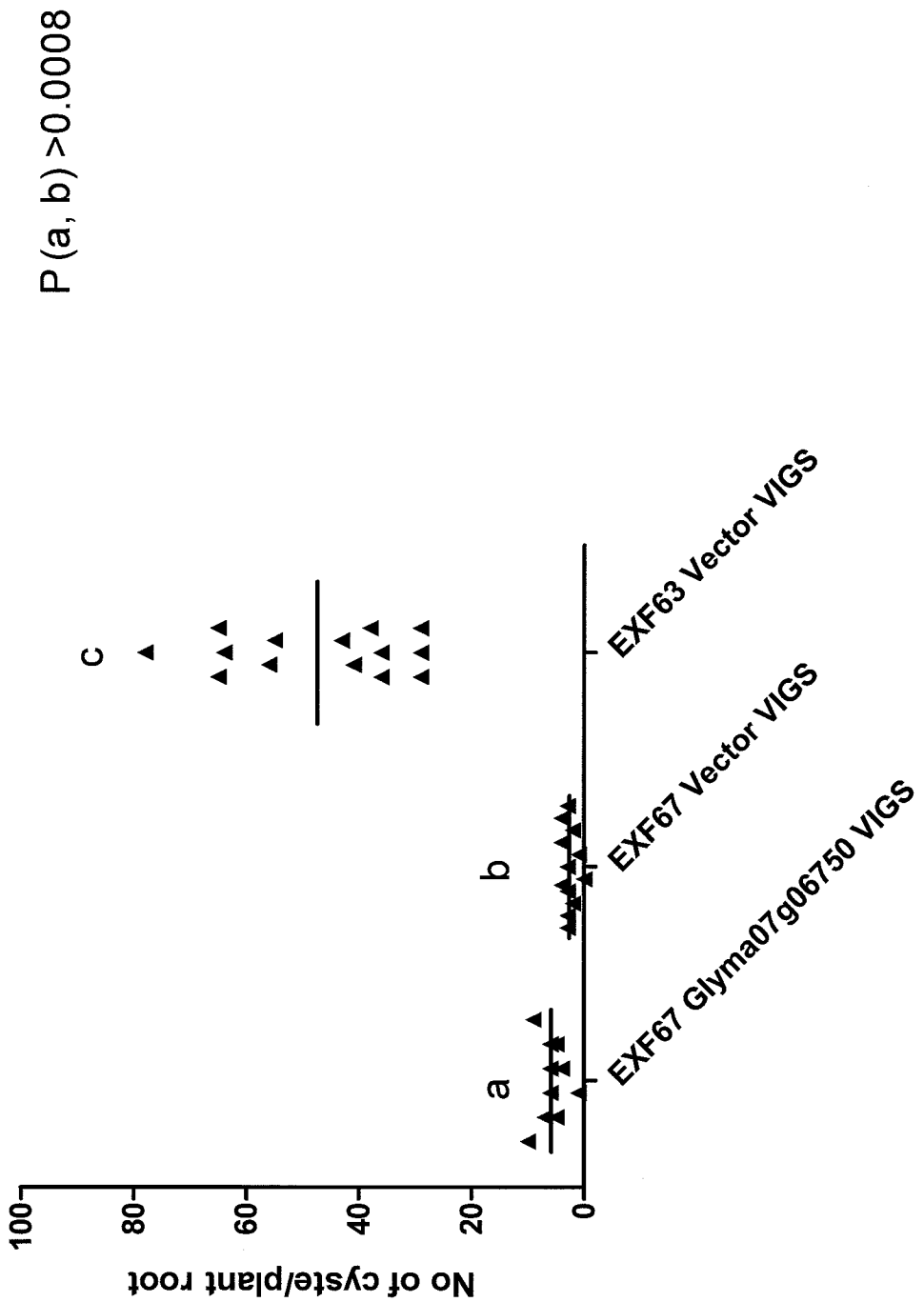

P (a, b) >0.0008
No of cyste/plant root
100
80
60
40
20
0
a
b
c
EXF67 Glyma07g06750 VIGS
EXF67 Vector VIGS
EXF63 Vector VIGS

FIG. 13

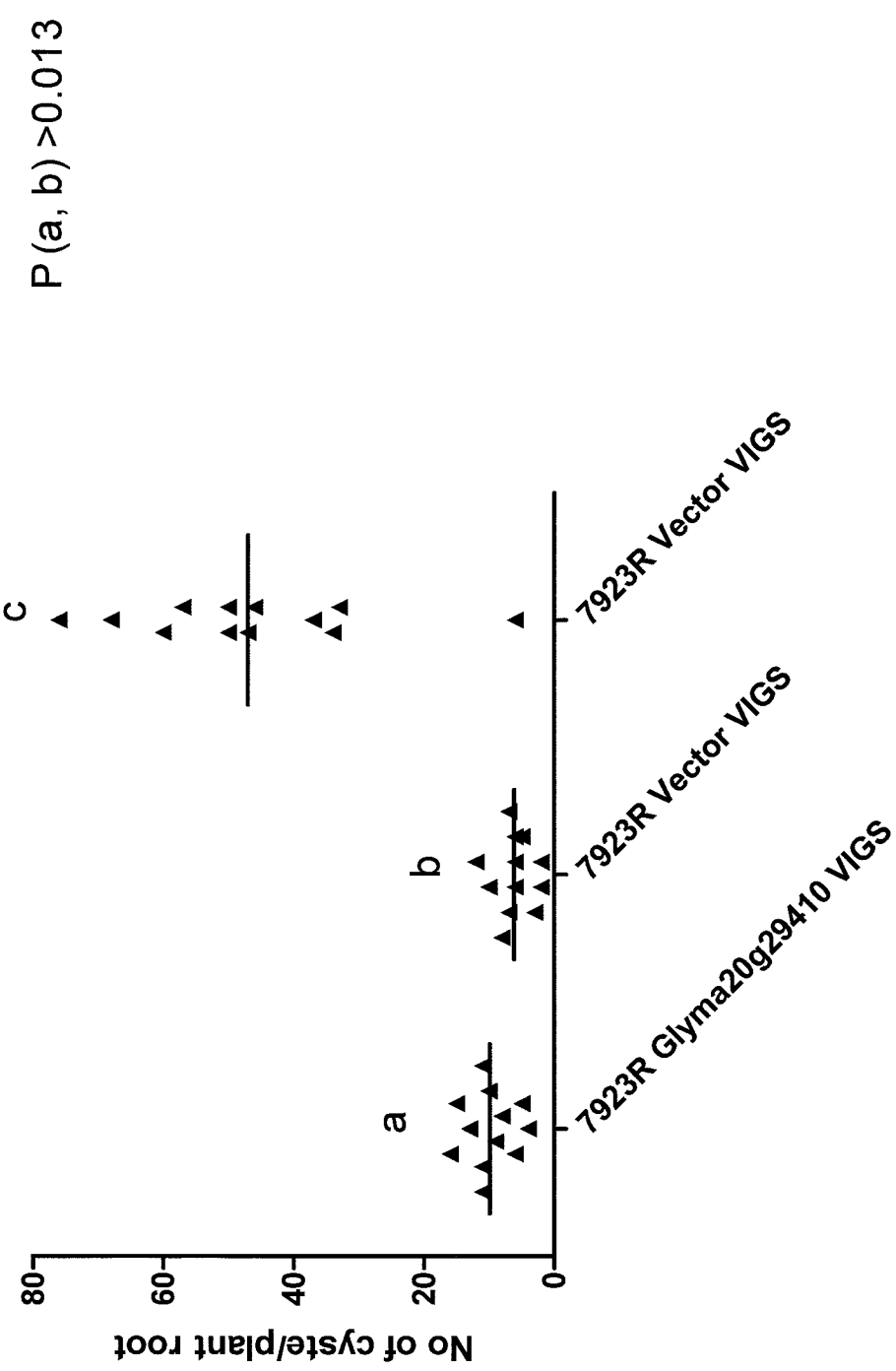

P (a, b) >0.013
No of cyste/plant root
80
60
40
20
0
a
b
c
7923R Glyma20g29410 VIGS
7923R Vector VIGS
7923R Vector VIGS

FIG. 14

GmBag6A and 6B Gene Information

GmBag6A sequence from 7923R

cDNA (3387 NT)(SEQ ID No. 162)

ATGAAGCTTGATCCATCCAAACCACCCTTTTCTTATGACCAACATTGGCCCTATGCCGGCAATTTTGGGCACCCTA
CTTCCCCACATTTCTGCTGTGGCCACAACAACTTCCCTTGTCATTATAGCTACATGCCTTCATATCCTCATGCCCC
TTCTCCAATGTACTATTCTGGAACTTGTCCTTCATATAGTGAACCATATTTTGTTCGTTATTCCCCACAACCACAT
TATACCATGGAGCTGCCTAGGTATGAAAATGACAAATGCATGCCCCGAGAGCTTCATTGTTCTGGTTCTGCTAATC
ATCCATGCAACCAAAAGGAAGGTAGAAGTGTGAAGATTGAAGAGCATGAACTGGATGGTGGAAAGAAAGAGAATGA
TGCTTTGGTCCCAATTCAGCTCAAGAATTATCCATATCCCTTAGTTTGGATTCCACAGGAGTACACAAGTAACAAA
CAGCTGAAAAATCCTAGTACAATGGAAGTTCGTGAACAAAACAAGCCTTCTAGTCTTGAGAATTCTAATGTTGATG
CACAGCCAACACAGGAGCCTATAGTATGGAATGGATGGCTTCCCTTCAATATAAAGGGTGCCCGGAACATGATTCA
CGATGGATATGGAACAAGAAACCAGAAACAGGAGTCTGGCAATAATAGAGGGGAATCTGAAAATGGAAAAATAGAC
CAGAAACATCAAAGTGAACAGAAGAGGTCAGAATTCCCATTCCCTATCTTCTGGTTGCCTTATTACAATAAGCAGG
AGGAGAGTGGAGAGACTAAGAACCAGGAGAAAAACATTTCTTCACCAAAAATTGTTGAGGAGGTACCCCATACATT
CAAATTTGTTCCAGTGAAGTCTCATGTTGATGAAGGTGGTAGGAACGGAACCGGATCAAATCAAGCTGATCAATCC
ACAAATACAAATGCTTCTTCGGATGCTGTAGAGAAAGTGAATAATGCCAGAAGCATACCTGTGAAGCAGATAGAAT
CCCACGAAGGAAAAAATGTTTCTCTCGATCAAATGGAAGAGAATGTGACACAAAAGGACTCTTGCACTGGGGACAA
AAAGAGACAATCTACATCTTCACCTAAAGGATCCAAGTTACCTCCGGTTTGTCTGAGAGTTGATCCACTACCAAGG
AAGAAAAATGGCCTCGGGAGTTCGAGTTCGAGGTCCCCAAGTCCACCTTCATCAAAAGGGAATTCCCAAGCTACAA
CTGGTGAAACATTCAAGACTCCTGTGAGTGGCACACGTGACAAGGCTCAGCCAAATTTGAATCATCAGAATGCTCC
AAACACCAGTGAGAAAGTTAAACCAAAGGAGAACACCATTCCGGTGTCAGAATGCATGACTAATGAAAACAAGGGT
GTTGACTGTAGGGATGGATGTCAGAGCCAGATGAAAGTAAACATACCCAGTAAAGGTCTGAAAGGGGCAAGGGAAA
CATGTCCAGATGATGATGACTATAAGACTGAAGATAAAAAGGCAGAGAAAGGAGCAGAAAATATGATGGAGGAAAC
TACTGAATCAAGGGAAGAGAAGGATTCAAGCACACGAACTGATGCAGGTCGAAAAGATGGAAGAGTTTTGTCAGAT
GCAGATGCTGCTGTTTTGATACAAGCTGCATATCGCAGTTATCTAGTTAGAAAATGGGAACCGTTGAAGACGTTGA
AGCAGATAGATGAAGTCAGGAAGGAGGTGACTCGTGTTCAAGGCCGTGTTCAAGCTTTTGAGAGATCTCCCGAACT
TCAAAATGATGACAAACAAAAAATTGCAATTGGAGAGACCATAATGAGACTCCTGCTGAAGTTGGATACTATACTG
GGTTTGCATCCAAGTTTCAGGGAGATCAGAAAATCCTTGGCTAGGGAGCTCATAATCTTGCAAGAAAGGCTTGATT
CTATAATGGCCAAGAAACCTCAGCAGCAGATGCCGGATGTTCAGGAACACGTTGAAATCACTCCAATGAACATGCA
GAGTGAAGAACATGTGCAAAAGCAGCAAGAAGAAAAGGTTGCTGTACCAGAGGATTCAGCTGAAGGCACTAGGGAT
GATGTAAAAGGTCCTTGTGCTAATGATGGTGGAAGTGAATCTCAGTCACCAGTTGATCCTCCATCAATTGAGGGAG
CAGAGTCTGTTGCACTTCCAAATGGCTCAGATAATGAGGACACCAGCCAAGTGGTTACATCTGATGCATTGAATTC
TTCAAGTGATCTGTCTGAGAGTGACAAAATGGCTGTGGAATCCGAAGCTAAATCAGAAGCGAAAGACAATCCGATT
GCGGAAGACATTCCCATTGAGGTTGATAAATTGGACAAGACTGTTTGGGAAGAATTGCCTGTGGGAGTTATTGATG
AAGATATCAATGATGTTAGTATTGAGAAGGAAGAACATGATGATGTTAGATCGGGAAGTCTCCCAGCCATGGTGAA
TGATTCGGCACAAGAAGGATTAAATTCAGAGAGCTATGCAATGATGGAACTGCCATTGGGATTACATGAGGAGCAT
GAAAGGGACAATGAAATGAATATTTCTAATGGAGAAACACGGTCTGAGAATGAGATATTTATTGAGGAGCTTCCTG
TGGGACTGCACGATGAAGATACAACAATATCTAAAGATAAGAGGGATGGTCAAGCTAAGCCTAAAACATATAAAGA
GGTTCGACTAGCTCAAGAAGGGGAATGCAATGCAGATGAGGAAACAAGTTCTTCCACAGATGACACTGCCAACGAA
ACTCAACTAGAGCAACAGCAGAAGCTGAAAGAGCAAGAAGAGGTGCATTCTTCTAGGGAATCAGATGGCTGGGTAA
AAATTGAGTACCCGGAAGAAGGTGAACTCAATGGTGATGCACCAATGGATATAAGAGTTGAGTGCAAGTCAGGTGA
GGAAGCTGGAACTGATACTAAGTTGCTTCCTTTAACAACACAAGTCAGTGATAATGAACCAGAAAATGAAGATGTA
TTCTCAGAAGCAAATTATGTAAATAACAAATTAACCGAGCCAATGGAGTTTGTACCTTCCAATGACACACAGAAGG
AGGAGACACCAGAGATGGTTGCTGAAGAGGCAATTATCCCTGATGATAAAGACACAGAAAATTTGGCCAAAGAGAA
AACTGAAGTATCTGCAGAACCACCACCTGCACTGCAAGACCGAGGGTTAAACGGTGACTCGAAGTTATTAGAAGAG
AATGAGAAGTTAAGGGAGATGATGAAGAAGTTGCTTGAAGCCGGGAATGAACAGTTAAGCGTGATATCAGATTTGA
CTGTCAGAGTGAAGGACTTGGAGAAGAAATTAGCCAGGAGAAGGAGTAAGAGAGTGAAGACAAAACAGTATAGACC
CGCAGCTTCCAAAATGTCTACCCATGAAATGAAATCCTCCTAA

FIG. 15

Protein (1129 AA) (SEQ ID No. 163)

MKLDPSKPPFSYDQHWPYAGNFGHPTSPHFCCGHNNFPCHYSYMPSYPHAPSPMYYSGTCPSYSEPYFVRYSPQPH
YTMELPRYENDKCMPRELHCSGSANHPCNQKEGRSVKIEEHELDGGKKENDALVPIQLKNYPYPLVWIPQEYTSNK
QLKNPSTMEVREQNKPSSLENSNVDAQPTQEPIVWNGWLPFNIKGARNMIHDGYGTRNQKQESGNNRGESENGKID
QKHQSEQKRSEFPFPIFWLPYYNKQEESGETKNQEKNISSPKIVEEVPHTFKFVPVKSHVDEGGRNGTGSNQADQS
TNTNASSDAVEKVNNARSIPVKQIESHEGKNVSLDQMEENVTQKDSCTGDKKRQSTSSPKGSKLPPVCLRVDPLPR
KKNGLGSSSSRSPSPPSSKGNSQATTGETFKTPVSGTRDKAQPNLNHQNAPNTSEKVKPKENTIPVSECMTNENKG
VDCRDGCQSQMKVNIPSKGLKGARETCPDDDDYKTEDKKAEKGAENMMEETTESREEKDSSTRTDAGRKDGRVLSD
ADAAVLIQAAYRSYLVRKWEPLKTLKQIDEVRKEVTRVQGRVQAFERSPELQNDDKQKIAIGETIMRLLLKLDTIL
GLHPSFREIRKSLARELIILQERLDSIMAKKPQQQMPDVQEHVEITPMNMQSEEHVQKQQEEKVAVPEDSAEGTRD
DVKGPCANDGGSESQSPVDPPSIEGAESVALPNGSDNEDTSQVVTSDALNSSSDLSESDKMAVESEAKSEAKDNPI
AEDIPIEVDKLDKTVWEELPVGVIDEDINDVSIEKEEHDDVRSGSLPAMVNDSAQEGLNSESYAMMELPLGLHEEH
ERDNEMNISNGETRSENEIFIEELPVGLHDEDTTISKDKRDGQAKPKTYKEVRLAQEGECNADEETSSSTDDTANE
TQLEQQQKLKEQEEVHSSRESDGWVKIEYPEEGELNGDAPMDIRVECKSGEEAGTDTKLLPLTTQVSDNEPENEDV
FSEANYVNNKLTEPMEFVPSNDTQKEETPEMVAEEAIIPDDKDTENLAKEKTEVSAEPPPALQDRGLNGDSKLLEE
NEKLREMMKKLLEAGNEQLSVISDLTVRVKDLEKKLARRRSKRVKTKQYRPAASKMSTHEMKSS*

Based on W82 and Phytozome

GmBag6A: Glyma07g06750 (Based on W82 and Phytozome)

gDNA (3858 NT)(SEQ ID No. 164)

ATGAAGCTTGATCCATCCAAACCACCCTTTTCTTATGACCAACATTGGCCCTATGCCGGCAATTTTGGGCACCCTA
CTTCCCCACATTTCTGCTGTGGCCACAACAACTTCCCTTGTCATTATAGCTACATGCCTTCATATCCTCATGCCCC
TTCTCCAATGTACTATTCTGGAACTTGTCCTTCATATAGTGAACCATATTTTGTTCCTTATTCCCCACAACCACAT
TATACCATGGAGCTGCCTAGGTATGAAAATGACAAATGCATGCCCCGAGAGCTTCATTGTTCTGGTTCTGCTAATC
ATCCATGCAACCAAAAGGAAGGTAGAAGTGTGAAGATTGAAGAGCATGAACTGGATGGTGGAAAGAAAGAGAATGA
TGCTTTGGTCCCAATTCAGCTCAAGAATTATCCATATCCCTTAGTTTGGATTCCACAGGAGTACACAAGTAACAAA
CAGCTGAAAAATCCTAGTACAATGGAAGTTCGTGAACAAAACAAGCCTTCTAGTCTTGAGAATTCTAATGTTGATG
CGCAGCCAACACAGGAGCCTATAGTATGGAATGGATGGCTTCCCTTCAATATAAAGGGTGCCCGGAACATGATTCA
CGATGGATATGGAACAAGAAACCAGAAACAGGAGTCTGGCAATAATAGAGGGGAATCTGAAAATGGAAAAATAGAC
CAGAAACATCAAAGTGAACAGAAGAGGTCAGAATTCCCATTCCCTATCTTCTGGTTGCCTTATTACAATAAGCAGG
AGGAGAGTGGAGAGACTAAGAACCAGGAGAAAAACATTTCTTCACCAAAAATTGTTGAGGAGGTACCCCATACATT
CAAATTTGTTCCAGTGAAGTCTCATGTTGATGAAGGTGGTAGGAACAGAACCGGATCAAATCAAGCTGATCAATCC
ACAAATACAAATGCTTCTTCGGATGCTGTAGAGAAAGTGAATAATGCCAGAAGCATACCTGTGAAGCAGATAGAAT
CCCACGAAGGAAAAAAATGTTTCTCTCGATCAAATGGAAGAGAATGTGACACAAAAGGACTCTTGCACTGGGGACAA
AAAGAGACAATCTACATCTTCACCTAAAGGATCCAAGTTACCTCCGGTTTGTCTGAGAGTTGATCCACTACCAAGG
AAGAAAAATGGCCACGGGAGTTCGAGTTCGAGGTCCCCAAGTCCACCTTCATCAAAAGGGAATTCCCAAGCTACAA
CTGGTGAAACATTCAAGACTCCTGTGAGTGGCACACATGACAAGGCTCAGCCAAATTTGAATCATCAGAATGCTCC
AAACACCAGTGAGAAAGTTAAACCAAAGGAGAACACCATTCCGGTGTCAGAATGCATGACTAATGAAAACAAGGGT
GTTGACTGTAGGGATGGATGTCAGAGCCAGATGAAAGTAAACATACCCAGTAAAGGTCTGAAAGGGGCAAGGGAAA
CATGTCCAGATGATGATGACTATAAGACTGAAGATAAAAAGGCAGAGAAAGGAGCAGAAAATATGATGGAGGAAAC
TACTGAATCAAGGGAAGAGAAGGATTCAAGCACACGAACTGATGCGGGTCGAAAAGATGGAAGAGTTTTGTCAGAT
GCAGATGCTGCTGTTTTGATACAAGCTGCATATCGCAGTTATCTAGTTAGAAAATGGGAACCGTTGAAGAAGTTGA
AGCAGATAGATGAAGTCAGGAAGGAGGTGACTCGTGTTCAAGGCCGTGTTCAAGCTTTTGAGAGATCTCCCGAACT
TCAAAATGATGACAAACAAAAAATTGCAATTGAAGAGACCATAATGAAACTCCTGCTGAAGTTGGATACTATACTG
GTACTTGTCAGATTACCCTATTTACTTATGATGAAATCTGTAGTCCACTTTTTGATATTCTTTGATTCATTGACAA
TAAACATTACACAGAATATTAAAGTGTTTTGATTTTGTATTTTACGGTCAAAATTCTCCTACGTATTTATTAGCTA
ATCTTAATTAATTGAGATTCATGACATCTGTGATAGAGTTTTGATATTATCTTATGATCTTATCTTGGGATATCAT
GGCATCAGTGACATAGTTTTGATATTTCTTCATTTTATGGATATGGAAAATGAATTTTTTTAAGCTATTAAAAAGC
ACTCTGCTTAACCGCATTTGAACTAAGATGAGCATTTTGATCTACACAGATAAGTGTTGTTTGTGGTTTGTCTGAG
CTATTAGTCATTGATCATCTTAATTATGATGGTTAATATAGTTCTCCTTTTTTGTTCAATGTTATGCTTACATTCG
TTCAATCTATATCAGGGTTTGCATCCAAGTTTCAGGGAGATCAGAAAATCCTTGGCTAGGGAGCTCATAATCTTGC
AAGAAAGGCTTGATTCTATAATGGCCAAGAAACCTCAGCAGCAGATGCCGGATGTTCAGGAACACGTTGAAATCAC
TCCAATGAACATGCAGAGTGAAGAACATGTGCAAAAGCAGCAAGAAGAAAAGGTTGCTGTACCAGAGGATTCAGCT
GAAGGCACTAGGGATGATGTAAAAGGTCCTTGTGCTAATGATGGTGGAAGTGAATCTCAGTCACCAGTTGATCCTC
CATCAAATGAGGGAGCAGAGTCTGTTGCACTTCCAAATGGCTCAGATAATGAGGACACCAGCCAAGTGGTTACATC

FIG. 15 cont

TGATGCATTGAATTCTTCAAGTGATCTGTCTGAGAGTGACAAAATGGCTGTGGAATCCGAAGCTAAATCAGAAGTG
AAAGACAATCCGATTGCGGAAGACATTCCCATTGAGGTTGATAAATTGGACAAGACTGTTTGGGAAGAATTGCCTG
TGGGAGTTATTGATGAAGATATCAATGATGTTAGTATTGAGAAGGAAGAACATGATGATATTAGATCGGGAAGTCT
CCCAGCCATGGTGAATGATTCGGCACAAGAAGGATTAAATTCAGAGAGCTATGCAATGATGGAACTGCCATTGGGA
TTACATGAGGAGCATGAAAGGGACAATGAAATGAATATTTCTAATGGAGAAACACGGTCTGAGAATGAGATATTTA
TTGAGGAGCTTCCTGTGGGACTGCACGATGAAGATACAACAATATCTAAAGATAAGAGGGATGGTCAAGCTAAGCC
TAAAACATATAAAGAGGTTCGACTAGCTCAAGAAGGGGAATGCAATGCAGATGAGGAAACAAGTTCTTCCACAGAT
GACACTGCCAACGAAACTCAACTAGAGCAACAGCAGAAGCTGAAAGAGCAAGAAGAGGTGCATTATTCTAGGGAAT
CAGATGGCTGGGTAAAAATTGAGTACCCGGAAGAAGGTGAACTCAATGGTGATGCACCAATGGATATAAGAGTTGA
GTGCAAGTCAGGTGAGGAAGCTGGAACTGATACTAAGTTGCTTCCTTTAACAACACAAGTCAGTGATAATGAACCA
GAAAATGAAGATGTATTCTCAGAAGCAAATTATGTAAATAACAAATTAACCGAGCCAATGGAGTTTGTACCTTCCA
ATGACACACAGAAGGAGGAGACACCAGAGATGGTTGCTGAAGAGGCAATTATCCCTGATGATAAAGACACAGAAAA
TTTGGCCAAAGAGAAAACTGAAGTATCTGCAGAACCACCACCTGCATTGCAAGACCGAGGGTTAAACGGTGACTCG
AAGTTATTAGAAGAGAATGAGAAGTTAAGGGAGATGATGAAGAAGTTGCTTGAAGCCGGGAATGAACAGTTAAGCG
TGATATCAGATTTGACTGTCAGAGTGAAGGACTTGGAGAAGAAATTAGCCAGGAGAAGGAGTAAGAGAGTGAAGAC
AAAACAGTATAGACCCGCAGCTTCCAAAATGTCTACCCATGAAATGAAATCCTCCTAA

cDNA (2991 NT)(SEQ ID No. 165)

ATGAAGCTTGATCCATCCAAACCACCCTTTTCTTATGACCAACATTGGCCCTATGCCGGCAATTTTGGGCACCCTA
CTTCCCCACATTTCTGCTGTGGCCACAACAACTTCCCTTGTCATTATAGCTACATGCCTTCATATCCTCATGCCCC
TTCTCCAATGTATGAAAATGACAAATGCATGCCCCGAGAGCTTCATTGTTCTGGTTCTGCTAATCATCCATGCAAC
CAAAAGGAAGGTAGAAGTGTGAAGATTGAAGAGCATGAACTGGATGGTGGAAAGAAAGAGAATGATGCTTTGGTCC
CAATTCAGCTCAAGAATTATCCATATCCCTTAGTTTGGATTCCACAGGAGTACACAAGTAACAAACAGCTGAAAAA
TCCTAGTACAATGGAAGTTCGTGAACAAAACAAGCCTTCTAGTCTTGAGAATTCTAATGTTGATGCGCAGCCAACA
CAGGAGCCTATAGTATGGAATGGATGGCTTCCCTTCAATATAAAGGGTGCCCGGAACATGATTCACGATGGATATG
GAACAAGAAACCAGAAACAGGAGTCTGGCAATAATAGAGGGGAATCTGAAAATGGAAAAATAGACCAGAAACATCA
AAGTGAACAGAAGAGGTCAGAATTCCCATTCCCTATCTTCTGGTTGCCTTATTACAATAAGCAGGAGGAGAGTGGA
GAGACTAAGAACCAGGAGAAAAACATTTCTTCACCAAAAATTGTTGAGGAGGTACCCCATACATTCAAATTTGTTC
CAGTGAAGTCTCATGTTGATGAAGGTGAGAAAGTGAATAATGCCAGAAGCATACCTGTGAAGCAGATAGAATCCCA
CGAAGGAAAAAATGTTTCTCTCGATCAAATGGAAGAGAATGTGACACAAAAGGACTCTTGCACTGGGGACAAAAAG
AGACAATCTACATCTTCACCTAAAGGATCCAAGTTACCTCCGGTTTGTCTGAGAGTTGATCCACTACCAAGGAAGA
AAAATGGCCACGGGAGTTCGAGTTCGAGGTCCCCAAGTCCACCTTCATCAAAAGGGAATTCCCAAGCTACAACTGG
TGAAACATTCAAGACTCCTAGCCAGATGAAAGTAAACATACCCAGTAAAGGTCTGAAAGGGGCAAGGGAAACATGT
CCAGATGATGATGACTATAAGACTGAAGATAAAAAGGCAGAGAAAGGAGCAGAAAATATGATGGAGGAAACTACTG
AATCAAGGGAAGAGAAGGATTCAAGCACACGAACTGATGCGGGTCGAAAAGATGGAAGAGTTTTGTCAGATGCAGA
TGCTGCTGTTTTGATACAAGCTGCATATCGCAGTTATCTAGTTAGAAAATGGGAACCGTTGAAGAAGTTGAAGCAG
ATAGATGAAGTCAGGAAGGAGGTGACTCGTGTTCAAGGCCGTGTTCAAGCTTTTGAGAGATCTCCCGAACTTCAAA
ATGATGACAAACAAAAAATTGCAATTGAAGAGACCATAATGAAACTCCTGCTGAAGTTGGATACTATACTGGGTTT
GCATCCAAGTTTCAGGGAGATCAGAAAATCCTTGGCTAGGGAGCTCATAATCTTGCAAGAAAGGCTTGATTCTATA
ATGGCCAAGAAACCTCAGCAGCAGATGCCGGATAGTGAAGAACATGTGCAAAAGCAGCAAGAAGAAAAGGTTGCTG
TACCAGAGGATTCAGCTGAAGGCACTAGGGATGATGTAAAAGGTCCTTGTGCTAATGATGGTGGAAGTGAATCTCA
GTCACCAGTTGATCCTCCATCAAATGAGGGAGCAGAGTCTGTTGCACTTCCAAATGGCTCAGATAATGAGGACACC
AGCCAAGTGGTTACATCTGATGCATTGAATTCTTCAAGTGATCTGTCTGAGAGTGACAAAATGGCTGTGGAATCCG
AAGCTAAATCAGAAGTGAAAGACAATCCGATTGCGGAAGACATTCCCATTGAGGTTGATAAATTGGACAAGACTGT
TTGGGAAGAATTGCCTGTGGGAGTTATTGATGAAGATATCAATGATGTTACCATGGTGAATGATTCGGCACAAGAA
GGATTAAATTCAGAGAGCTATGCAATGATGGAACTGCCATTGGGATTACATGAGGAGCATGAAAGGGACAATGAAA
TGAATATTTCTAATGGAGAAACACGGTCTGAGAATGAGATATTTATTGAGGAGCTTCCTGTGGGACTGCACGATGA
AGATACAACAATATCTAAAGATAAGAGGGATGGTCAAGCTAAGCCTAAAACATATAAAGAGGTTCGACTAGCTCAA
GAAGGGGAATGCAATGCAGATGAGGAAACAAGTTCTTCCACAGATGACACTGCCAACGAAACTCAACTAGAGCAAC
AGCAGAAGCTGAAAGAGCAAGAAGAGGTGCATTATTCTAGGGAATCAGATGGCTGGGTAAAAATTGAGTACCCGGA
AGAAGGTGAACTCAATGGTGATGCACCAATGGATATAAGAGTTGAGTGCAAGTCAGGTGAGGAAGCTGGAACTGAT
ACTAAGTTGCTTCCTTTAACAACACAAGTCAGTGATAATGAACCAGAAAATGAAGATGTATTCTCAGAAGCAAATT
ATGTAAATAACAAATTAACCGAGCCAATGGAGTTTGTACCTTCCAATGACACACAGAAGGAGGAGACACCAGAGAT
GGTTGCTGAAGAGGCAATTATCCCTGATGATAAAGACACAGAAAATTTGGCCAAAGAGAAAACTGAAGTATCTGCA
GAACCACCACCTGCATTGCAAGACCGAGGGTTAAACGGTGACTCGAAGTTATTAGAAGAGAATGAGAAGTTAAGGG
AGATGATGAAGAAGTTGCTTGAAGCCGGGAATGAACAGTTAAGCGTGATATCAGATTTGACTGTCAGAGTGAAGGA
CTTGGAGAAGAAATTAGCCAGGAGAAGGAGTAAGAGAGTGAAGACAAAACAGTATAGACCCGCAGCTTCCAAAATG
TCTACCCATGAAATGAAATCCTCCTAA

FIG. 15 (cont)

Protein (996 AA)(SEQ ID No. 166)

MKLDPSKPPFSYDQHWPYAGNFGHPTSPHFCCGHNNFPCHYSYMPSYPHAPSPMYENDKC
MPRELHCSGSANHPCNQKEGRSVKIEEHELDGGKKENDALVPIQLKNYPYPLVWIPQEYT
SNKQLKNPSTMEVREQNKPSSLENSNVDAQPTQEPIVWNGWLPFNIKGARNMIHDGYGTR
NQKQESGNNRGESENGKIDQKHQSEQKRSEFPFPIFWLPYYNKQEESGETKNQEKNISSP
KIVEEVPHTFKFVPVKSHVDEGEKVNNARSIPVKQIESHEGKNVSLDQMEENVTQKDSCT
GDKKRQSTSSPKGSKLPPVCLRVDPLPRKKNGHGSSSSRSPSPPSSKGNSQATTGETFKT
PSQMKVNIPSKGLKGARETCPDDDDYKTEDKKAEKGAENMMEETTESREEKDSSTRTDAG
RKDGRVLSDADAAVLIQAAYRSYLVRKWEPLKKLKQIDEVRKEVTRVQGRVQAFERSPEL
QNDDKQKIAIEETIMKLLLKLDTILGLHPSFREIRKSLARELIILQERLDSIMAKKPQQQ
MPDSEEHVQKQQEEKVAVPEDSAEGTRDDVKGPCANDGGSESQSPVDPPSNEGAESVALP
NGSDNEDTSQVVTSDALNSSSDLSESDKMAVESEAKSEVKDNPIAEDIPIEVDKLDKTVW
EELPVGVIDEDINDVTMVNDSAQEGLNSESYAMMELPLGLHEEHERDNEMNISNGETRSE
NEIFIEELPVGLHDEDTTISKDKRDGQAKPKTYKEVRLAQEGECNADEETSSSTDDTANE
TQLEQQQKLKEQEEVHYSRESDGWVKIEYPEEGELNGDAPMDIRVECKSGEEAGTDTKLL
PLTTQVSDNEPENEDVFSEANYVNNKLTEPMEFVPSNDTQKEETPEMVAEEAIIPDDKDT
ENLAKEKTEVSAEPPPALQDRGLNGDSKLLEENEKLREMMKKLLEAGNEQLSVISDLTVR
VKDLEKKLARRRSKRVKTKQYRPAASKMSTHEMKSS*

GmBag6B:Glyma16g03320

gDNA (3850 NT) (SEQ ID No. 167)

ATGAAGCTTGATCCATCCAAACCACCCTTTTCTTATGACCAACATTGGCCATATGCTGGCAACTTAGGGCACCCTA
TTCCCCCACATTTCTGCTGTGGCCACAACAACTTCCCTTGTCATTATAGCTACATGCCTTCAAATCCTCATGCTCC
TTCTCCAATGTACTATTCTGGAACTTGTCCTTCATATAGTGAACCATATTTTGTTCCTTATTCACTACAACCACAT
TATACTATGGATCTGCCCAGGTATGAATATGACAAATTTATGCCCCGAGAGCATCATTGTTGTGGCTGTCCTAATC
ATCCATGCAGCCAAAAGGAAGGTAGAAGTGTGAAGATTGAAGAACATGAACCTGATGGTGGAAAGAAAGTGAATGA
TGCTTTGGCTCCAATCCAGCTCAAGAATTACCCATATCCCTTAGTTTGGATTCCACAGGAGTACACAAGTAACAAA
CAGCTGAAGAATCCTAATACAATGGAAGTTGGGGAACAAGACAAGTCTCCTCGTTTTGAGAATTCTAATGCTGATG
CACAGCCAGCACAGGATCCTAGAGTATGGAATGGATGGCTTCCCTTCAATATAAAGGGTGCCCCGAACGTGATTCA
CGATGGATATGGAACAAGAAACCAGAAACAGGAGTCTGGCAATAATAGAGGGGAATCTGAAGATGGAAAAGTGGAC
CAGAAACATCAAAGTGAACAGAATAGGTCAGAATTTCCATTCCCTATCTTCTGGTTGCCTTATTACAATAAGCCGG
AGGAGGGTGGTGAGACTAAAAACCAGGAGAAGAATACTTTTTCACCAAAAGTTATTGAGGAGGTACCCCATACATT
AAAATCTGTTCCAGCAAAGTCTCATGTTGATGAAGGTGGTAGGAATAGGAACGGAACCGGATCAAATCAAGCTGAA
TACACAAATACAAATGCTGCGGATGTTGTAGAGAAAGTGACTAATGCCAGAAGCATACCTGTGAAGCAAATGGAAG
AGAATGTGACAA
AAAAGGACTCCTGCACTGGTGACAAAAAGAGGCAATCTGCATCTTCACCAAAAGGATCCAAGTTACCTCCGGTTTG
TCTGAGAGTTGATCCACTACCAAGGAAGAAAAATGGCAACGGGAGTTTGAGTTTGAGTTCAAGGTCGAGGTCCCCA
AGTCCACCTTCATCAAAAGGGCAATCCCAAGCTACAACTGGTGAAACATTCAAGACTCCTGTGAGTGGAACACATG
ATAAAGGTCAGCCAAATTCGGATCATCAGAATGGGCCAACAAACACCAGTGAGAGAGTTAAACCAAAGGAGAAGAC
CATTCCGGTGTCAGAAATATGCAAGACTAATGAAAACCAGGGTGTTGAGTGTAGGGATGGATGTCAAAGCCAGATA
AATGTAAACATACCTAGTGAAGATCCGAAAGGGACAAGGGGAACATGTCCAGATGGTGATGACTTCAAGACTGAAG
ATAAAAAGGCAGATAAAGGGGCAGATAAAATGAGGGAGAAAACTGCAGAACTAAAAGATGTGAAGGATTCAAGCGC
ACAAACTGACGCAGGTCAAAAGGAGGTAAGAGTTCTGTCAGATGCAGATGCTTCTGTTTTGATACAAGCTGCATAT
CGCGGACATCAAGTTAGAAAATGGGAAGCATTGAAGAAGTTGAAGCAGATAGATGAAGTCAGGAAGGAGGTGACTG
ATGTTCGAGACCGTGTTCAAGCCTTTGAGAGATCTTCTGATCTTCAAAATGATGACAAGCAAAAAATTGCAATTGG
AGAGACCATTATGAGACTCCTGCTGAAGTTGGATACTATACAGGTACATGTCAGATAACCGTTACTTAAAATGAAA
TCAGTAATGCATTTTTTATATTTTTGATTCATTGACAATAAACATTACATGGAATATGAAATTGTTTTGCTTTTGT
ACTTCAGTAAAAAGATTCCATTTTTATTTATTAGGTAATCTTAATTGAGATATCATGACATCAGAGTATAGTTTTG
ATATTAGCTAATCTTAGTTAGATATCATGACATCTGTAATATAGTTTTGATATTTCTTCATTCATGGATATGGGTA
GTGACTTTTACATGCCTAATTACCTGTTACAACTTACAAGCACTCCGCATTTGAACAAAGATTAGCATTTTGATCT
GCACAGATATGAATGTTGGTTGTGGTTTGTCTGAGCTATTAATCGTTGTTCATCTTATTTATGATTTTTTTTAATT
TAGTTATACTTCCTTGTTCAATGTTATGCTTACATATCGCTCAACATTTATATCAGGGTTTGCATCCAAGTATCAG
GGAGATCAGAAAATCCTTGGCTAGAGAGCTGACAATTTTGCAAGAAAGGCTTGATTATATAATGGCCAAGAAACCT
CAGCAGCAGGTGCAGGATTTCAATGTTCAGGAACATGTTGAAGTCACTCCAATGAATATGCAGAATGAAGAACAAG
TTGCTGTACCAGCAGACTCAGCTGAAGGCATTAGTGATGATGCAAAAGGTCCTTGTGCTAATGTTGGTGGAAGTGA

FIG. 15 (cont)

ATCCCAGTCACCAGTTGATCCTACATCAGATAAGGGAGCAGAGTCTGTTGTGCTTCCGAATGGCTCGCATAATGAT
GATACTAGCCAAGTGGTTACAGATGACACATTGAATTCTACAAGTGATCTGTCTGAGACGGACAAAATGGCTGTGG
AACCCGAAGCTAAATCAGAAGCGAAAGACATTCCCACTGCGGAAGACATTCTAAACGAGGTTGATAAATTGTATAA
GACTGTTTTGGAAGAATTTCCTGTGGGAGTTGTTGATGAAGATATCGATGATGTTATTATTGAGAAGGAAGAACAA
GATGATGTTAGTTCTGGAAGTCTCCAAGCCATGGTGGATGATTCGGCACGAGATGGATTAGAGAACCATGCAATGA
TGGAAATGCCAGTGGGATTATTTGATGAGGATGAAAGGAACAATGAAATGAATATTTCTAAAGAGGGAACACAAAC
TGAGAATGACATATTTATTGAGCTTCCTATGGGACTACTTGATGAGGATACAGCAATATCTGAAGTTAAGAAGCAC
GATCGAGCTAAGCCTAAAACATATAAAGAGGTTCTACTAGCTCAAGAAGGGGAATGCAATGCAGATGAGAAAACAA
GTTCTTCAACAGATAACACTGCGAACGAAACTCAACTAGGGCAACAGCAGAAGCTGGAAGAGCAAAAAGAGGTGCA
TTCTTCTGGGGAATGCTGGGTGAAAATCAAGTACCAGGAAGAAGGTGAACTCAATGGTGATGCACCAATAGATATA
AGAACTGAGTGTAAGTCAGGTGAGGAAGTTGAAAATGATACTAAGTTACCTCCTTTGACAACACAAGTCAATGATC
ATGGACCAGGAAATGAAGATGTATGCTCAAAAGCAAATGATGTAAATAACATATCGCCCGAGGCCAAGCCAATGGA
GTTAGTACCTTCCAATGAAACACAGAAGGAGGAGGAACCAGAAGAGAAGATTGCACTTAGGGAGACACAAGAGATG
GTTACTGAGCAGGCAATTGGTGCTGATGATAAAGACCCAGAAGCTTTGTTTAAAGAGAAAACCGAACTATCTTCGC
CTCCACCTGCATTACAAGAACCTGAGCTGCCTGCGGCCATGCGCGATGGAGAGTTCAACGGTGACACGAAGTTACT
GGAAGAGAATGAGAATTTAAGGGAGATGATGAAGAAGTTGCTTGAAGCCGGGAATGAACAGTTAAGCGTGATAACA
AATTTGACTGGCAGAGTGAAGGACTTGGAGAAGAAACTAGCCAGGAGCAGGAGTAAGAGAGTGAAGACAAAACGGT
ATAGACCAGCAACTTCCAAAATGTCTTGCATGAAATGA

cDNA (2715 NT) (SEQ ID No. 168)

ATGAAGCTTGATCCATCCAAACCACCCTTTTCTTATGACCAACATTGGCCATATGCTGGCAACTTAGGGCACCCTA
TTCCCCCACATTTCTGCTGTGGCCACAACAACTTCCCTTGTCATTATAGCTACATGCCTTCAAATCCTCATGCTCC
TTCTCCAATGTATGAATATGACAAATTTATGCCCCGAGAGCATCATTGTTGTGGCTGTCCTAATCATCCATGCAGC
CAAAAGGAAGGTAGAAGTGTGAAGATTGAAGAACATGAACCTGATGGTGGAAAGAAAGTGAATGATGCTTTGGCTC
CAATCCAGCTCAAGAATTACCCATATCCCTTAGTTTGGATTCCACAGGAGTACACAAGTAACAAACAGCTGAAGAA
TCCTAATACAATGGAAGTTGGGGAACAAGACAAGTCTCCTCGTTTTGAGAATTCTAATGCTGATGCACAGCCAGCA
CAGGATCCTAGAGTATGGAATGGATGGCTTCCCTTCAATATAAAGGGTGCCCCGAACGTGATTCACGATGGATATG
GAACAAGAAACCAGAAACAGGAGTCTGGCAATAATAGAGGGGAATCTGAAGATGGAAAAGTGGACCAGAAACATCA
AAGTGAACAGAATAGGTCAGAATTTCCATTCCCTATCTTCTGGTTGCCTTATTACAATAAGCCGGAGGAGGGTGGT
GAGACTAAAAACCAGGAGAAGAATACTTTTTCACCAAAAGTTATTGAGGAGGTACCCCATACATTAAAATCTGTTC
CAGCAAAGTCTCATGTTGATGAAGGTGGTAGGAATAGGAACGGAACCGGATCAAATCAAGCTGAATACACAAATAC
AAATGCTGCGGATGTTGTAGAGAAAGTGACTAATGCCAGAAGCATACCTGTGAAGCAAATGGAAGAGAATGTGACA
AAAAAGGACTCCTGCACTGGTGACAAAAAGAGGCAATCTGCATCTTCACCAAAAGGATCCAAGTTACCTCCGGTTT
GTCTGAGAGTTGATCCACTACCAAGGAAGAAAAATGGCAACGGGAGTTTGAGTTTGAGTTCAAGGTCGAGGTCCCC
AAGTCCACCTTCATCAAAAGGGCAATCCCAAGCTACAACTGGTGAAACATTCAAGACTCCTGTGAGTGGAACACAT
GATAAAGATCCGAAAGGGACAAGGGGAACATGTCCAGATGGTGATGACTTCAAGACTGAAGATAAAAAGGCAGATA
AAGGGGCAGATAAAATGAGGGAGAAAACTGCAGAACTAAAAGATGTGAAGGATTCAAGCGCACAAACTGACGCAGG
TCAAAAGGAGGTAAGAGTTCTGTCAGATGCAGATGCTTCTGTTTTGATACAAGCTGCATATCGCGGACATCAAGTT
AGAAAATGGGAAGCATTGAAGAAGTTGAAGCAGATAGATGAAGTCAGGAAGGAGGTGACTGATGTTCGAGACCGTG
TTCAAGCCTTTGAGAGATCTTCTGATCTTCAAAATGATGACAAGCAAAAAATTGCAATTGGAGAGACCATTATGAG
ACTCCTGCTGAAGTTGGATACTATACAGGGTTTGCATCCAAGTATCAGGGAGATCAGAAAATCCTTGGCTAGAGAG
CTGACAATTTTGCAAGAAAGGCTTGATTATATAATGGCCAAGAAACCTCAGCAGCAGGTGCAGGATTTCAATGTTC
AGGAACATGTTGAAGTCACTCCAATGAATATGCAGAATGAAGAACAAGTTGCTGTACCAGCAGACTCAGCTGAAGG
CATTAGTGATGATGCAAAAGGTCCTTGTGCTAATGTTGGTGGAAGTGAATCCCAGTCACCAGTTGATCCTACATCA
GATAAGGGAGCAGAGTCTGTTGTGCTTCCGAATGGCTCGCATAATGATGATACTAGCCAAGTGGTTACAGATGACA
CATTGAATTCTACAAGTGATCTGTCTGAGACGGACAAAATGGCTGTGGAACCCGAAGCTAAATCAGAAGCGAAAGA
CATTCCCACTGCGGAAGACATTCTAAACGAGGTTGATAAATTGTATAAGACTGTTTTGGAAGAATTTCCTGTGGGA
GTTGTTGATGAAGATATCGATGATGTTATTATTGAGAAGGAAGAACAAGATGATGTTAGTTCTGGAACAATATCTG
AAGTTAAGAAGCACGATCGAGCTAAGCCTAAAACATATAAAGAGGTTCTACTAGCTCAAGAAGGGGAATGCAATGC
AGATGAGAAAACAAGTTCTTCAACAGATAACACTGCGAACGAAACTCAACTAGGGCAACAGCAGAAGCTGGAAGAG
CAAAAAGAGGTGCATTCTTCTGGGGAATGCTGGGTGAAAATCAAGTACCAGGAAGAAGGTGAGGAAGTTGAAAATG
ATACTAAGTTACCTCCTTTGACAACACAAGCAATTGGTGCTGATGATAAAGACCCAGAAGCTTTGTTTAAAGAGAA
AACCGAACTATCTTCGCCTCCACCTGCATTACAAGAACCTGAGCTGCCTGCGGCCATGCGCGATGGAGAGTTCAAC
GGTGACACGAAGTTACTGGAAGAGAATGAGAATTTAAGGGAGATGATGAAGAAGTTGCTTGAAGCCGGGAATGAAC
AGTTAAGCGTGATAACAAATTTGACTGGCAGAGTGAAGGACTTGGAGAAGAAACTAGCCAGGAGCAGGAGTAAGAG
AGTGAAGACAAAACGGTATAGACCAGCAACTTCCAAAATGTCTTGCATGAAATGA

FIG. 15 (cont)

Protein (904 AA) (SEQ ID No. 169)

MKLDPSKPPFSYDQHWPYAGNLGHPIPPHFCCGHNNFPCHYSYMPSNPHAPSPMYEYDKFMPREHHCCGCPNHPCS
QKEGRSVKIEEHEPDGGKKVNDALAPIQLKNYPYPLVWIPQEYTSNKQLKNPNTMEVGEQDKSPRFENSNADAQPA
QDPRVWNGWLPFNIKGAPNVIHDGYGTRNQKQESGNNRGESEDGKVDQKHQSEQNRSEFPFPIFWLPYYNKPEEGG
ETKNQEKNTFSPKVIEEVPHTLKSVPAKSHVDEGGRNRNGTGSNQAEYTNTNAADVVEKVTNARSIPVKQMEENVT
KKDSCTGDKKRQSASSPKGSKLPPVCLRVDPLPRKKNGNGSLSLSSRSRSPSPPSSKGQSQATTGETFKTPVSGTH
DKDPKGTRGTCPDGDDFKTEDKKADKGADKMREKTAELKDVKDSSAQTDAGQKEVRVLSDADASVLIQAAYRGHQV
RKWEALKKLKQIDEVRKEVTDVRDRVQAFERSSDLQNDDKQKIAIGETIMRLLLKLDTIQGLHPSIREIRKSLARE
LTILQERLDYIMAKKPQQQVQDFNVQEHVEVTPMNMQNEEQVAVPADSAEGISDDAKGPCANVGGSESQSPVDPTS
DKGAESVVLPNGSHNDDTSQVVTDDTLNSTSDLSETDKMAVEPEAKSEAKDIPTAEDILNEVDKLYKTVLEEFPVG
VVDEDIDDVIIEKEEQDDVSSGTISEVKKHDRAKPKTYKEVLLAQEGECNADEKTSSSTDNTANETQLGQQQKLEE
QKEVHSSGECWVKIKYQEEGEEVENDTKLPPLTTQAIGADDKDPEALFKEKTELSSPPPALQEPELPAAMRDGEFN
GDTKLLEENENLREMMKKLLEAGNEQLSVITNLTGRVKDLEKKLARSRSKRVKTKRYRPATSKMSCMK*

FIG. 15 (cont)