Plant Gene 4 (2015) 1–9

Contents lists available at ScienceDirect

Plant Gene

journal homepage: www.elsevier.com/locate/plant-gene

A genome-wide survey of glycolytic genes in diploid Asian cotton (*Gossypium arboreum*)

Bing Zhang

Institute of Botany, Jiangsu Province and Chinese Academy of Sciences, Nanjing 210014, China
State Key Laboratory of Cotton Biology, Institute of Cotton Research, Chinese Academy of Agricultural Sciences, Anyang 455000, China

ARTICLE INFO

Article history:
Received 20 August 2014
Received in revised form 25 August 2015
Accepted 1 September 2015
Available online 15 September 2015

Keywords:
Asian cotton
Glycolysis
Genome-wide analysis
cis-Regulatory element
Hypoxia

ABSTRACT

Glycolysis is an indispensable biochemical pathway in both animals and plants and is required for almost all physiological processes. Previous studies have shown that glycolysis plays an important role in cotton fiber development process. However, the detailed mechanism by which glycolysis is regulated is still unclear. In this study, a total of 163 genes encoding different isoforms of 20 enzymes involved in the catalysis of glycolysis and fermentation were identified in diploid Asian cotton (*Gossypium arboreum*). These enzymes have unique subcellular localizations and phylogenies. Among these 163 genes, 125 were expressed in Asian cotton plants, but only 51 were highly expressed in elongating Asian cotton fibers. Cis-regulatory elements involved in phytohormone responses were identified in these 51 genes, suggesting that glycolysis might be regulated by phytohormones. Furthermore, the expression of fermentation-related genes and the wide distribution of cis-regulatory elements that promote anaerobic induction strongly suggested the involvement of anaerobic glycolysis in cotton fiber development, especially during the elongation process.

1. Introduction

As the central metabolic pathway in all organisms, and especially in non-photosynthetic cells, glycolysis supplies cells with energy and metabolic intermediates for the synthesis of amino acids, nucleic acids, fatty acids and diverse secondary metabolites (Bar-Even et al., 2012). Through the 10 steps of glycolysis, one molecule of glucose and two molecules of ADP and NAD^+ are converted into two molecules of pyruvate and two molecules of ATP and NADH. To compensate for the loss of NAD^+, pyruvate can be further metabolized, either oxidatively to CO_2 in the tricarboxylic acid (TCA) cycle via oxidative phosphorylation, or reductively to lactate and ethanol via fermentation reactions that regenerate NAD^+. Based on the availability of O_2, glycolysis, in combination with fermentation, can be either aerobic or anaerobic (Gatenby and Gillies, 2004). Although aerobic glycolysis generates energy less efficiently than the TCA cycle, many rapidly proliferating cells, such as *Saccharomyces cerevisiae* and cancer cells, use aerobic glycolysis as the major metabolic pathway for glucose, even in the presence of sufficient O_2 to support the TCA cycle with oxidative phosphorylation. These phenomena are called the Crabtree effect and the Warburg effect, respectively (Lunt and Vander Heiden, 2011).

Unlike non-photosynthetic organisms, green plants can acquire energy and metabolic intermediates through photosynthesis instead of glycolysis. In non-photosynthetic plant tissues, sucrose is often used instead of glucose as the principal substrate for glycolysis (Plaxton, 1996). In addition, glycolytic enzymes can exist in both the cytosol and the plastids, resulting in a complex cytosol–plastid transport system for glycolytic intermediates and more complex regulatory mechanisms for each reaction (Plaxton, 1996). In addition, the final glycolytic products in plant cells could be both pyruvate and malate; the former is catalyzed by pyruvate kinase (PK), as in animal cells, whereas the latter is converted from phosphoenolpyruvate (PEP) through a two-step reaction that is sequentially catalyzed by phosphoenolpyruvate carboxylase (PEPC) and malate dehydrogenase (MDH). As an added layer of complexity, malate can also be converted to pyruvate through a decarboxylic reaction catalyzed by malic enzyme (ME) (Fernie et al., 2004). For these reasons, although some key regulatory mechanisms of the glycolytic pathway have been elucidated in animal cells, the roles of glycolysis in the growth, development, stress responses and other physiological processes of plants are still poorly understood.

Gossypium arboreum, commonly known as Asian cotton or tree cotton, is one of the major cultivated cotton species worldwide, especially in tropical area of south Asia (Li et al., 2014). Along with the fiberless diploid cotton *Gossypium raimondii*, diploid *G. arboreum* is speculated as the putative progenitor species of the new world allotetraploid upland cotton *Gossypium hirsutum* and sealand cotton *Gossypium barbadense* (Hovav et al., 2008). Compared with two allotetraploid cotton species, diploid Asian cotton has shorter mature fiber, which inevitably restricts its commercial value. However, the development process

E-mail address: bingkeith@gmail.com.

http://dx.doi.org/10.1016/j.plgene.2015.09.001

of cotton fiber cell in Asian cotton, including initiation, elongation and secondary cell wall deposition, is similar to that of allotetraploid cotton, suggesting the genome-simplified diploid *G. arboreum* can be used as a model species for molecular mechanism research of cotton fiber development instead of the genome-complicated allotetraploid cotton.

In the past several years, the key roles of carbohydrate metabolism, and especially glycolysis, in cotton fiber development have gradually been recognized. The accumulation of malate and K^+ in the vacuole was suspected to increase turgor pressure, driving fiber elongation (Dhindsa et al., 1975). The expression and activity of PEPC, the enzyme responsible for malate synthesis, were shown to correlate with the fiber elongation rate (Li et al., 2010). Sucrose transported from the basal cells into the adjacent cotton fiber cells could be further catalyzed by either sucrose synthase (SUS) or invertase (INV), both of which were found to be indispensable for the proper initiation and elongation of cotton fibers (Ruan et al., 2003; Wang et al., 2010). Furthermore, comparative proteomic analyses of cotton fiber cells at five time points during the elongation process revealed that glycolysis was the most significantly regulated biochemical pathway (Zhang et al., 2013). Together, the results of these studies indicated that glycolysis is essential for cotton fiber development. However, the detailed molecular mechanisms by which glycolysis is regulated, especially at the transcriptional level, are still unclear. In this study, through a genome-wide survey of glycolytic genes in Asian cotton (*G. arboreum*) and analysis of their expression in elongating fiber cells, we systematically analyzed the transcriptional regulation of glycolysis in diploid cotton fibers and determined that anaerobic glycolysis might exist in the rapidly elongating cotton fiber cells.

2. Methods

2.1. Database search and sequence retrieval

To identify all the glycolytic genes in Asian cotton (*G. arboreum*), multiple database searches were performed. Firstly, the Arabidopsis glycolytic gene sequences were retrieved from the Arabidopsis information resource (http://www.arabidopsis.org) using ontologies/keywords search interface with glycolysis as keyword. Next, coding sequence of these Arabidopsis glycolytic genes were used as queries to perform repetitive local blast searches against the downloaded *G. arboreum* genome data (http://www.cgp.genomics.org.cn) (Li et al., 2014). All output genes with an E-value < 1.0 were collected and translated to proteins, which were further blast searched against protein sequence data repositories at the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/) to confirm the identity as glycolytic enzymes.

2.2. Prediction of subcellular localization and phylogenetic analysis

The subcellular localization of each glycolytic protein was predicted using TargetP (www.cbs.dtu.dk/services/TargetP) (Emanuelsson et al., 2000). For each glycolytic enzyme family, the amino acid sequences of its members were aligned using ClustalX (version 2.1) (Larkin et al., 2007). Phylogenetic trees were then constructed using MEGA software (version 5.05) with the Neighbor-Joining (NJ) method and 1000 bootstrap replicates (Tamura et al., 2011).

2.3. Dataset-based gene expression analysis

To determine the relative expression levels of the glycolytic genes in Asian cotton, all gene sequences were BLAST searched against the *G. arboreum* transcriptome assembly (NCBI Sequence Read Archive SRR952685 with 1,699,776 total reads, which was transformed to a fasta file by NCBI SRA Toolkit) with E-value = 0 (Guan et al., 2014). The expression level of each gene was then calculated based on the matched read numbers and gene lengths using the RPKM algorithm (Mortazavi et al., 2008). The expression levels of glycolytic genes in elongating fibers (10 days post anthesis) were obtained from a microarray dataset (NCBI Gene Expression Omnibus datasets GSE52432, contributed by Hande et al. at National Research Centre on Plant Biotechnology, New Delhi, India) generated using an Affymetrix cotton GeneChip Genome array that compared the transcriptome profiles of *G. arboreum* L. cv. and its fuzzy-lintless mutant, ANOI 1960, during fiber development and statistically analyzed as described (Benjamini and Hochberg, 1995).

2.4. Real-time RT-PCR analysis

The total RNA was extracted from the roots, leaves, stems, petals and 10 days post anthesis (DPA) fibers of *G. arboreum* L. cv. DPL971 using the RNAprep pure Plant kit (TIANGEN, Beijing, China). Each RNA sample was treated with DNase I after extraction to remove residual DNA. The qRT-PCR assays were performed using the SYBR Premix Ex Taq (TaKaRa, Dalian, China) and a MX3000P Real-Time PCR system (Stratagene, California, USA) with gene-specific primers (Table S1). A cotton ubiquitin gene (UBQ7, DQ116441) was used as a standard control. All experiments were repeated for 3 times. The relative gene expression level was calculated using the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001).

2.5. Promoter sequence analysis

The 2000 bp of upstream sequences before the translation initiation codon ATG of each Asian cotton glycolytic genes was selected as gene promoters. Cis-regulatory elements of each promoter sequences were predicted though searching in PlantCARE database (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/) (Lescot et al., 2002). Fischer's exact tests with an FDR threshold of ≤0.1% were performed to further retrieve the high-confidence cis-regulatory elements (Yazawa et al., 2013).

3. Results

3.1. Identification of glycolytic genes in G. arboreum

To identify the complete set of glycolytic genes in Asian cotton, the chromosomes and scaffold sequences of *G. arboreum* were locally BLAST searched using known glycolytic gene sequences from Arabidopsis as queries. As shown in Table S2, a total of 138 genes encoding glycolytic enzymes and 25 genes involved in lactate and ethanol fermentation were successfully identified. Almost all 163 genes were mapped to the 13 chromosomes of the *G. arboreum* genome; the exceptions were GaSUS7, GaSUS8, GaUGP3 (UDP-glucose pyrophosphorylase) and GaLDH3 (lactate dehydrogenase), which were found on the scaffold sequences. Members of the identified gene families correspond to enzymes that catalyze 20 different glycolytic and fermentation reactions, were named according to their chromosomal order (Fig. 1).

Several glycolytic genes are distributed on the chromosomes as gene clusters. For example, GaHXK3 and GaHXK4, which encode hexokinases; GaPFK9 and GaPFK10, which encode phosphofructokinases; GaGAP8 and GaGAP9, which encode glyceraldehyde-3-phosphate dehydrogenases; GaPGK2 and GaPGK3, which encode phosphoglycerate kinases; and PK16 and PK17, which encode pyruvate kinases, were all arranged in tandem (Fig. 1). Even more striking was alcohol dehydrogenase, an important enzyme in fermentation; GaADH5 and GaADH6, GaADH8 and GaADH9, GaADH12 and GaADH13, GaADH15, GaADH16 and GaADH17 were located in 4 clusters of tandem repeats on four chromosomes. (Fig. 1). Among these gene clusters, GaPK16 and GaPK17 shared 99.2% sequence similarity, strongly suggesting that these two genes might have arisen through a gene duplication event (Table S3).

It is well known that plant glycolytic enzymes are distributed in both the cytosol and in plastids. For example, in developing castor oil seeds, both the cytosol and plastids were reported to contain complete

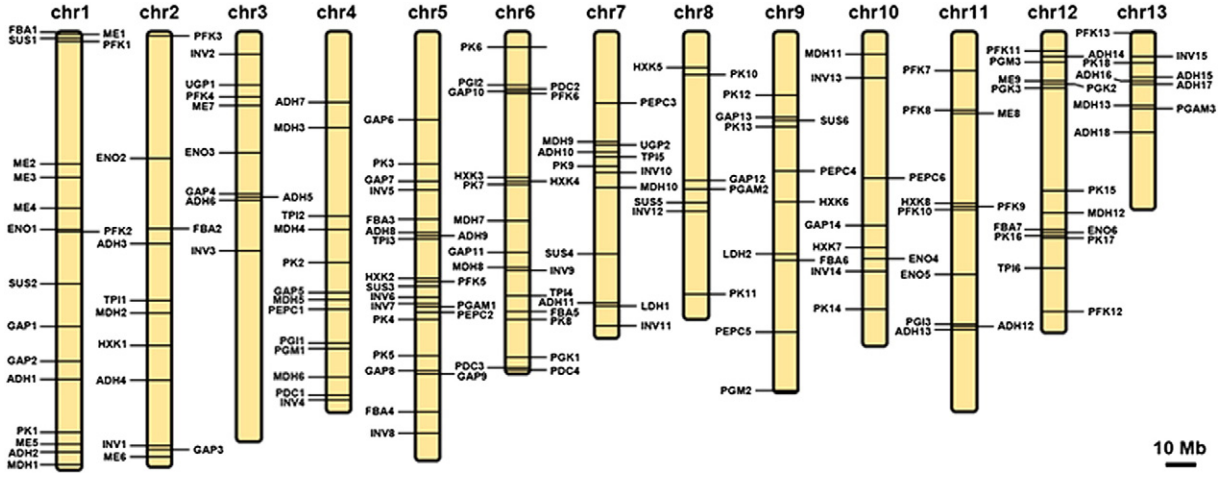

Fig. 1. Distributions of glycolytic genes in the Asian cotton genome. The physical locations of the 159 glycolytic genes on the 13 chromosomes of Asian cotton are indicated by the gene names. Chromosome names are indicated at the top. The scale bar represents 10 megabases (Mbs).

glycolytic pathways that can convert hexose phosphate to pyruvate (Miernyk and Dennis, 1982). To determine if this is also the case in cotton plants, the predicted subcellular localization of the glycolytic gene products was analyzed. As shown in Table S2, except for invertase, sucrose synthase, UDP-glucose pyrophosphorylase, phosphoglycerate mutase (PGAM) and phosphoenolpyruvate carboxylase, all glycolytic enzymes had at least one gene encoding a plastid-localized isoform. Notably, the products of the three phosphoglucomutase (PGM) genes were all found in the plastids, suggesting that the interconversion of glucose-1-phosphate and glucose-6-phosphate, which is catalyzed by PGM, should occur entirely in the plastids of Asian cotton cells. This finding was surprising because cytosolic isoforms of PGM exist in many other plants, such as Arabidopsis, maize and potato, and they play important roles in maintaining sink–source relationships (Fernie et al., 2007). Because developing cotton fibers (sink) constantly consume the sucrose synthesized by photosynthetic tissues (source), the deficiency of cytosolic PGM isoforms might be a unique property of Asian cotton.

3.2. Phylogenetic analyses of G. arboreum glycolytic genes

To determine the evolutionary relationship between glycolytic genes that encode the same enzymes, the amino acid sequences of the gene products were used to construct phylogenetic trees. All the gene families with 3 genes, including UGP, PGM, PGI (glucose-6-phosphate isomerase), PGK and PGAM, share a similar phylogenetic tree structure: 2 of the 3 genes are clustered in one clade and are separated from the remaining gene, suggesting that these gene families formed in a similar way (Fig. 2-D, E, F, K, L). As shown in Fig. 2-A, the 15 GaINV proteins clustered in 3 groups, in agreement with the known classification of invertase proteins into 3 types (cell wall, vacuolar and cytoplasmic) (Sturm, 1999). The 8 GaSUS proteins also clustered into 3 groups (Fig. 2-B), consistent with another genome-wide analysis showing that cotton sucrose synthase genes could be divided to 3 classes (Zou et al., 2013). The phylogenetic trees for the GaHXK and GaENO proteins were similar in that both of them contained 3 groups: group I contained most of the cytosolic isoforms, group II contained the plastid isoforms and group III contained only one cytosolic isoform (Fig. 2-C and M). As shown in Fig. 2-G, the 13 GaPFK proteins could be divided to 2 groups, both of which contain cytosolic and plastid isoforms. Sequence analysis further indicated that one of these groups comprised ATP-dependent phosphofructokinases and the other comprised pyrophosphate-dependent phosphofructokinases (Mustroph et al., 2013). The phylogenetic trees for GaFBA (fructose-bisphosphate aldolase) and GaTPI (triosephosphate isomerase) were also similar; both of them were composed of 2 groups, one containing the cytosolic isoforms and the other containing the plastid isoforms (Fig. 2-H and I). Similar tree structures were obtained from the phylogenetic analyses of 14 GaGAP and 18 GaPK proteins, except that the cytosolic and plastid isoforms were further divided into subgroups according to their sequence divergence (Fig. 2-J and N).

The GaMDH and GaME phylogenetic trees were more complex because some of the isoforms are located in the mitochondria (MDH and ME) and peroxisomes (MDH). As shown in Fig. 2-P, 13 GaMDH proteins clustered into 4 groups, each of which corresponded to one subcellular location. In contrast, for the GaME proteins, phylogenetic analysis indicated that isoforms with different subcellular localizations were clustered within the same groups (Fig. 2-Q), suggesting that the differences between signal peptide sequences were smaller than those between the functional domains. Phylogenetic analyses of enzymes involved in lactate and ethanol fermentation showed that LDH, which catalyzes lactate formation, formed 2 classes whose members were all located in the plastids (Fig. 2-S), whereas PDC (pyruvate decarboxylase) and ADH, which both catalyze ethanol fermentation, could be divided into 3 classes and were almost all located in the cytosol; the exception was ADH11, which was located in the plastids and was phylogenetically distant from the other ADH proteins (Fig. 2-R and T).

3.3. Differential expression of glycolytic genes in elongating Asian cotton fiber cells

Members of a gene family often have different expression patterns (Puranik et al., 2012). The expression of different glycolytic genes in whole Asian cotton plants was first analyzed via a BLAST search of high-throughput RNA sequencing data. As shown in Fig. 3 and Table S4, 125 of the 163 glycolytic genes were expressed in Asian cotton plants. Notably, enzymes catalyzing the key rate-limiting steps of glycolysis, such as INV, SUS, GAP, PFK and PK, all had more than 8 genes that were differentially expressed, implying that the regulation of glycolysis was extremely complex in Asian cotton plants. Furthermore, the expression levels of these 125 genes varied greatly, not only between the different enzymes, but also between different isoforms of the same enzyme. For example, the expression of PK, which is encoded by 18 GaPK genes, was approximately 13-fold of that of LDH, which is encoded by 3 GaLDH genes, and the expression of the GaSUS5 gene was 383-fold that of GaSUS1, which has similar catalytic activity (Table S4).

The expression of glycolytic genes in rapidly elongating Asian cotton fibers was then analyzed using microarray data comparing their expression in fiber-bearing ovules and the ovules of the fuzzy-lintless mutant. Among the 125 glycolytic genes expressed in Asian cotton plants, only

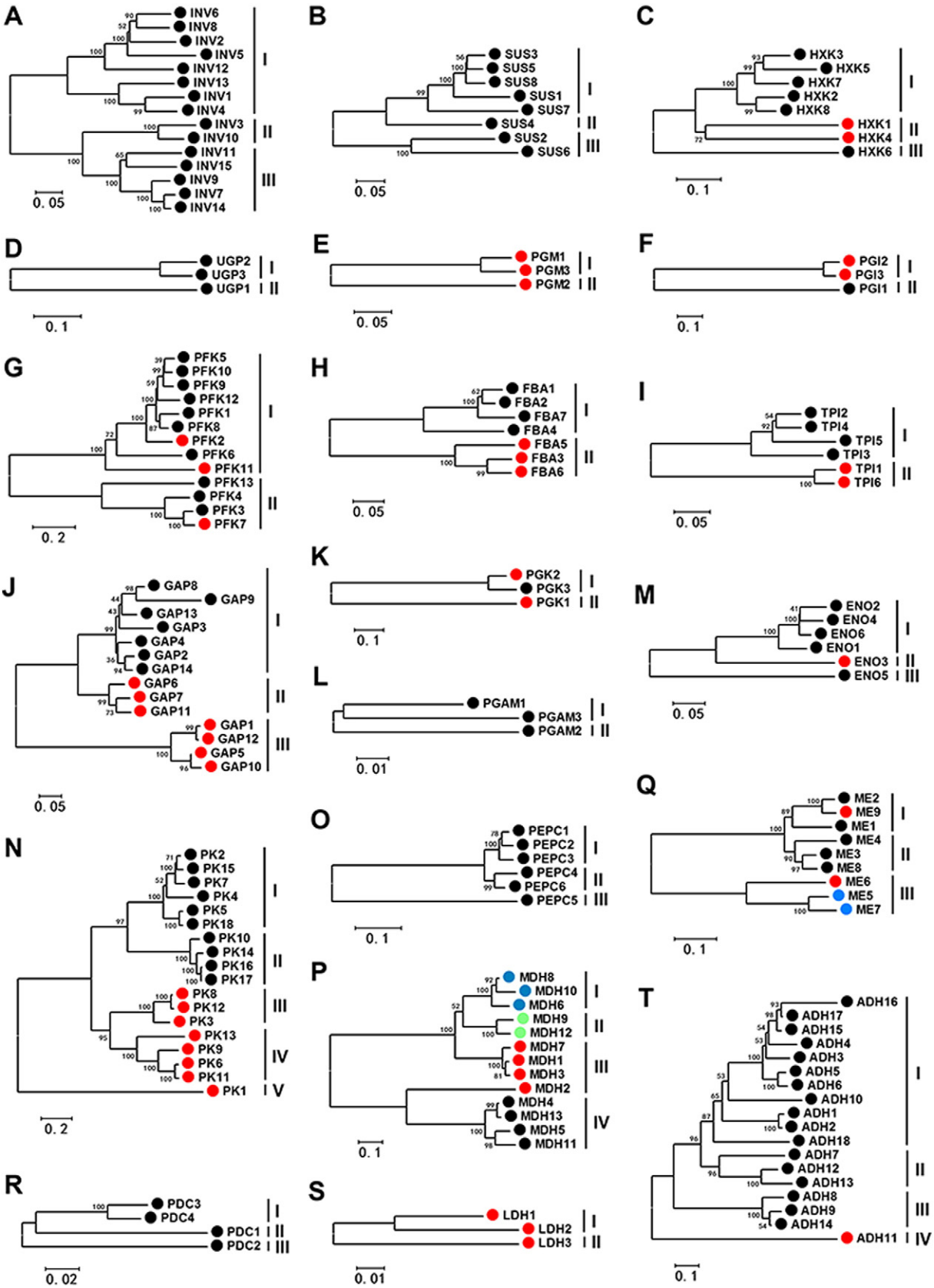

Fig. 2. Phylogenetic relationships of glycolytic proteins in Asian cotton. Phylogenetic trees were constructed using MEGA 5.05 based on the complete protein sequence alignment of the A) invertase, B) sucrose synthase, C) hexokinase, D) UDP-glucose pyrophosphorylase, E) phosphoglucomutase, F) glucose-6-phosphate isomerase, G) phosphofructokinase, H) fructose-bisphosphate aldolase, I) triosephosphate isomerase, J) glyceraldehyde-3-phosphate dehydrogenase, K) phosphoglycerate kinase, L) phosphoglycerate mutase, M) enolase, N) pyruvate kinase, O) phosphoenolpyruvate carboxylase, P) malate dehydrogenase, Q) malic enzyme, R) pyruvate decarboxylase, S) lactate dehydrogenase, T) alcohol dehydrogenase enzyme family. The numbers beside the branches indicate the bootstrap values that support the adjacent node. Black, red, blue and green circles indicate localization in the cytoplasm, plastids, mitochondria and peroxisomes, respectively.

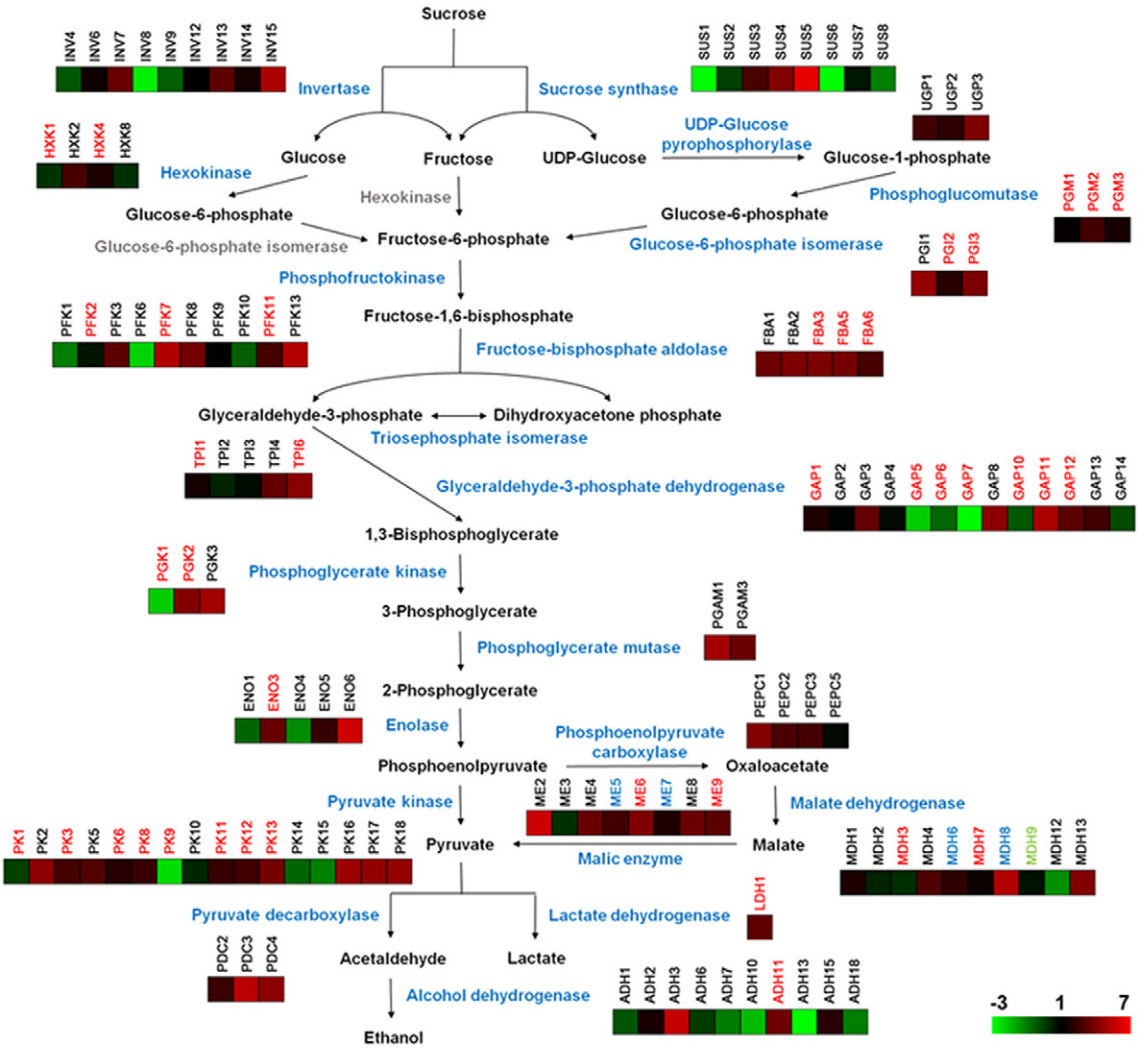

Fig. 3. Expression analysis of glycolytic genes in Asian cotton plants. The glycolysis pathway is shown as a schematic diagram. A heat map representing the log2 transformed RPKM gene expression values is displayed beside the names of the corresponding enzymes. Gene names in black, red, blue and green indicate subcellular localization in the cytoplasm, plastids, mitochondria and peroxisomes, respectively.

51 were expressed in elongating fibers (Fig. 4 and Table S5). Although most of the 51 genes were high-abundance genes, with abundant transcripts detected in whole cotton plants, 6 fiber-specific genes, GaSUS6, GaHXK1, GaGAP4, GaPK10, GaPK14 and GaADH18, were preferentially expressed in elongating cotton fibers (Figs. 3 and 4). These analysis results were supported by quantitative RT-PCR results that expression level of all the 6 genes in elongating fibers were at least 5-fold higher than that of other tissues (Fig. 5). The analysis also revealed an interesting phenomenon: none of the GaUGP and GaPGM genes were expressed in elongating cotton fibers, meaning that UDP-glucose derived from sucrose cannot be glycolytically consumed (Fig. 4). Because rapidly elongating cotton fibers must synthesize a large amount cellulose for their cell walls using UDP-glucose as the raw material (Ruan et al., 2003), using transcriptional regulation to prevent UDP-glucose from entering the glycolytic pathway could effectively promote cell wall expansion and fiber cell elongation.

As mentioned above, plant glycolytic enzymes have both cytosol- and plastid-localized isoforms. Among the 51 glycolytic genes expressed in elongating Asian cotton fibers, 16 genes encode enzyme isoforms that are localized to the plastids (Fig. 4). In particular, GaFBA3, GaFBA5 and GaLDH1, which encode two plastid-located isoforms of fructose-bisphosphate aldolase and one of lactate dehydrogenase, respectively, were the only expressed FBA and LDH genes, suggesting that fructose-1, 6-bisphosphate decomposition and lactate fermentation only occurred in the plastids of elongating cotton fiber cells. In contrast, sucrose breakdown, the transformation of 1, 3-bisphosphoglycerate to phosphoenolpyruvate and ethanol fermentation likely all occurred in the cytosol of Asian cotton fiber cells because the enzymes catalyzing these reactions, including GaINV7, GaSUS6, GaPGK3, GaPGAM3, GaENO6, GaPDC2 and GaADH18, are all located in the cytosol (Fig. 4).

3.4. Cis-regulatory elements in glycolytic genes expressed in elongating Asian cotton fiber cells

Cis-regulatory elements in gene promoters determine the specificity and level of gene expression (Todeschini et al., 2014). Different genes that are involved in a common biochemical or signaling pathway are often co-expressed because they possess similar cis-regulatory elements (Usadel et al., 2009). Accordingly, cis-regulatory element analysis of the 163 Asian cotton glycolytic genes using the PLACE database revealed that several well-characterized cis-regulatory elements were conserved in the 2000 bp up-stream promoter regions (Table S6). For example, a cis-acting regulatory element essential for anaerobic induction (ARE) could be identified in 136 glycolytic gene promoters (Table S6), suggesting that these genes might be induced by low oxygen levels. MYB transcription factors were well known to be involved in regulating cotton fiber initiation and elongation (Pu et al., 2008). Among

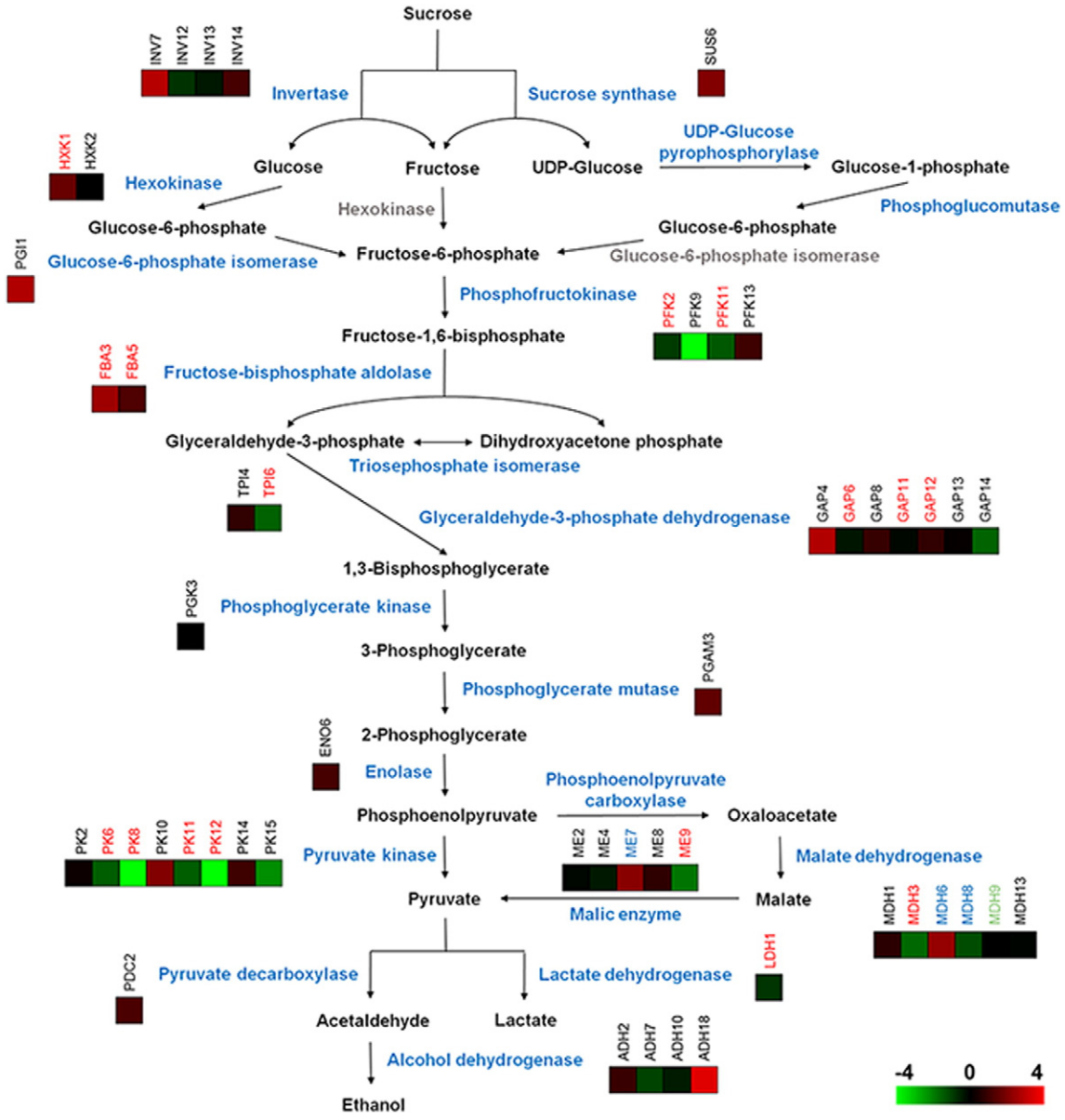

Fig. 4. Expression analysis of glycolytic genes in elongating Asian cotton fibers. The glycolysis pathway is shown as a schematic diagram. A heat map representing the log2 transformed microarray-derived gene expression values is displayed beside the names of the corresponding enzymes. Gene names in black, red, blue and green indicate subcellular localization in the cytoplasm, plastids, mitochondria and peroxisomes, respectively.

the 133 genes that have MYB transcription factor binding sites in their promoters, 43 were expressed in elongating cotton fibers (Table S6 and Fig. 6), suggesting that they might be down-stream MYB targets in Asian cotton fiber cells.

Phytohormones are important regulators of plant development. Cotton fiber development, and especially elongation, is known to be regulated by auxin, gibberellic acid (GA) and ethylene (Shi et al., 2006; Xiao et al., 2010; Zhang et al., 2011). Among the 51 glycolytic genes expressed in elongating Asian cotton fibers, 25, 36 and 30 genes were found to have cis-regulatory elements specific for auxin, GA and ethylene responses in their respective promoters (Table S6 and Fig. 6). Furthermore, 13 genes, including GaINV7, GaPFK9, GaGAP4, GaGAP11, GaGAP12, GaGAP14, GaENO6, GaPK15, GaMDH1, GaMDH13, GaME2, GaME8 and GaME9, were found to have all the 3 cis-regulatory elements in their promoters (Fig. 6), strongly suggesting that the activity of the glycolytic pathway in elongating Asian cotton fibers is synergistically regulated by different phytohormones. Interestingly, among these 13 genes, GaINV7 and GaGAP4 were the most highly expressed in their gene families, whereas GaPFK9, GaGAP14 and GaME9 were the lowest expressed members (Fig. 4), suggesting that the regulation of gene expression by these cis-regulatory elements could be either positive or negative.

4. Discussion

Carbohydrate metabolism, and especially glycolysis, lies at the center of plant life, providing cells with the energy and diverse metabolic intermediates required for development, stress responses and other physiological processes (Dennis and Miernyk, 1982). As one of the longest and most quickly elongating plant cells, the cotton fiber is the most important natural raw material in the textile industry (Mansoor and Paterson, 2012). Cotton fiber development, including initiation, elongation and secondary wall differentiation, determine the final length and cellulose content of the mature fiber, making it an interesting and valuable research subject. Previous studies revealed that, during fiber development, the content of intracellular carbohydrates such as soluble

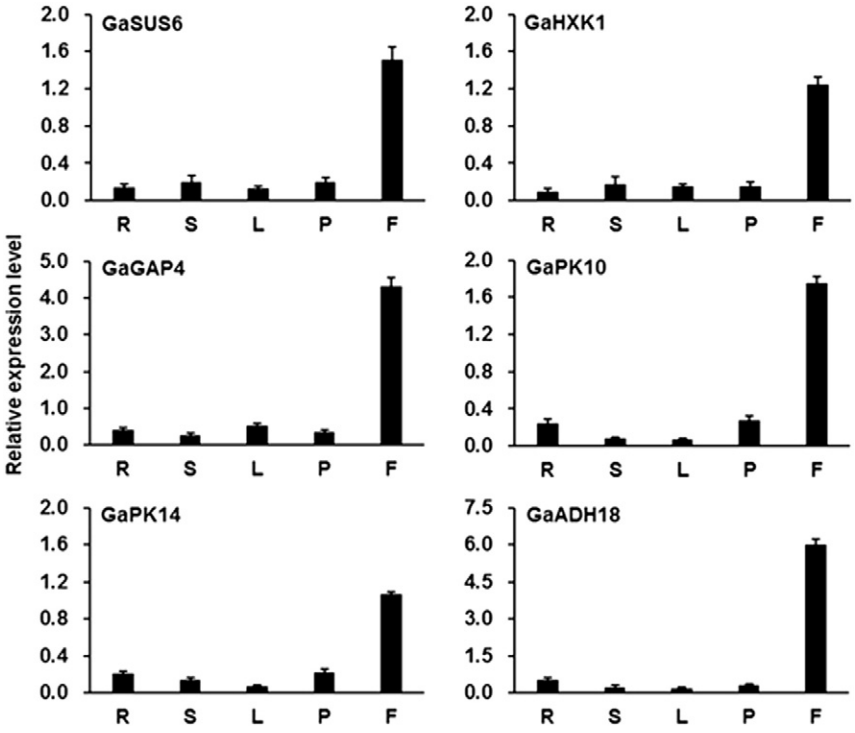

Fig. 5. qRT-PCR analysis of fiber-specific glycolytic genes in cotton tissues. The relative expression level of 6 fiber-specific glycolytic genes, GaSUS6, GaHXK1, GaGAP4, GaPK10, GaPK14 and GaADH18, in Asian cotton roots (R), stems (S), leaves (L), petals (P) and 10 DPA fibers (F) were determined using qRT-PCR. Error bars indicate standard deviation.

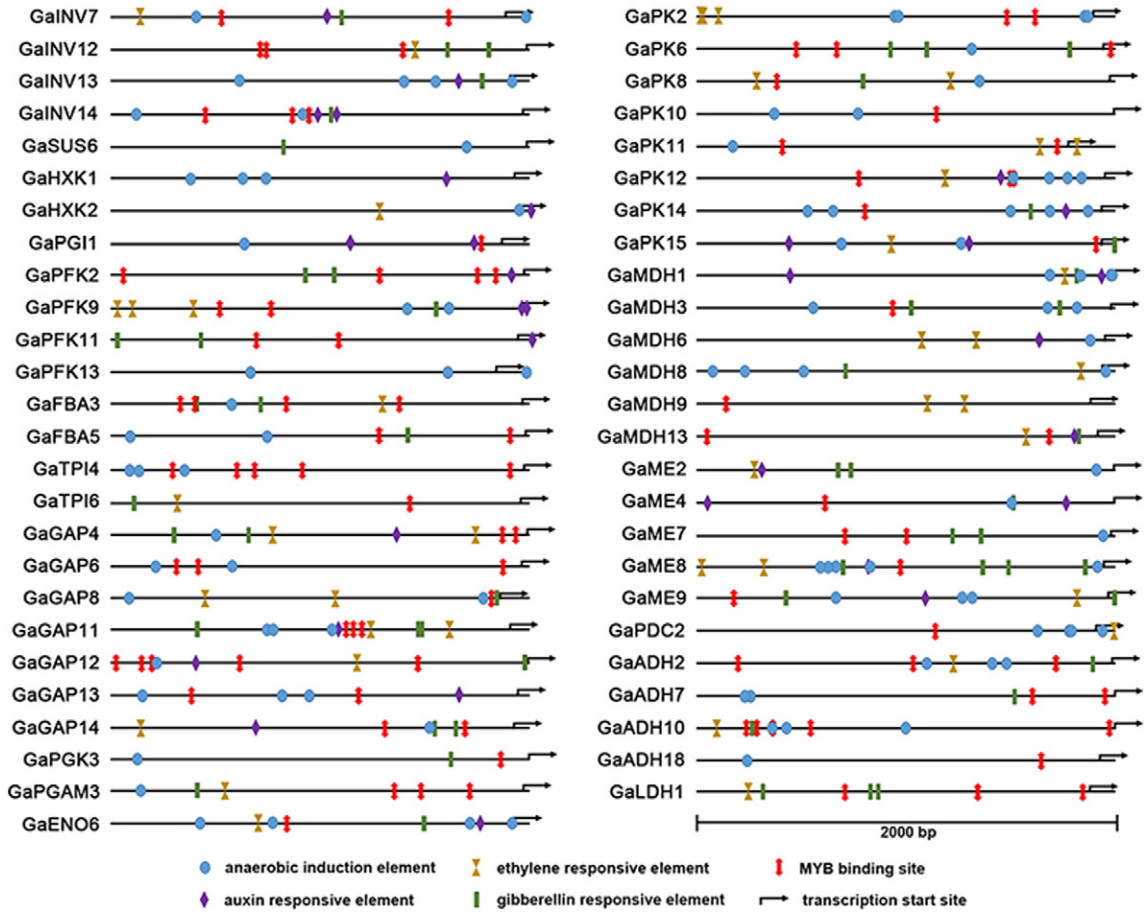

Fig. 6. Putative cis-regulatory elements in the promoters of glycolytic genes with enriched expression in Asian cotton fiber cells. Relative positions of cis-regulatory elements and transcription start sites are shown on the line representing the 2000 bp up-stream region of Asian cotton glycolytic gene promoters. Only cis-regulatory elements required for auxin, GA and ethylene responses, anaerobic induction and MYB transcription factor binding are shown.

sugars and organic acids was dynamic and correlated with specific developmental stages, suggesting that the regulation of carbohydrate metabolism is critical for proper fiber development (Gou et al., 2007; Naoumkina et al., 2013). In this study, through a genome-wide survey of glycolytic genes in diploid Asian cotton and analysis of the expression of these genes in elongating Asian cotton fibers, 51 essential glycolytic genes encoding enzyme isoforms that catalyze glycolysis and fermentation were successfully identified (Fig. 4). The great heterogeneity in the expression of these 51 genes implied that different glycolytic steps and different enzyme isoforms were both independently and precisely regulated in diploid cotton fibers. This finding was in agreement with previous reports that isoforms of many plant glycolytic enzymes were differentially regulated (Claeyssen and Rivoal, 2007; Miller et al., 1998; Mustroph et al., 2013; Sturm, 1999; Van der Straeten et al., 1991). In animal cells, the activity of different glycolytic enzyme isoforms correlates with the efficiency of glycolysis (Christofk et al., 2008; Zancan et al., 2010). In cotton, differential expression of glycolytic enzyme isoforms enables the fiber cell to precisely regulate the glycolytic pathway, more efficiently utilizing assimilated carbon sources.

The ARE cis-regulatory element was first identified in the maize ADH1 gene promoter (Walker et al., 1987) and was later was found in the promoters of many Arabidopsis genes that are induced under low-oxygen stress (Klok et al., 2002). Microarray analysis revealed that genes involved in glycolysis and fermentation are among the most strongly induced under low-oxygen stress (Liu et al., 2005). In agreement with these reports, the ARE cis-regulatory element was found in as many as 90% (45/51) of the promoters of glycolytic genes expressed in elongating Asian cotton fibers (Fig. 6). This result, together with the finding that fermentation-related genes such as GaPDC2, GaADH2, GaADH7, GaADH10, GaADH18 and GaLDH1 were also expressed (Figs. 4 and 5), strongly suggested that cotton fiber development, and especially elongation, might occur in hypoxic conditions. This hypothesis is reasonable because cotton fibers, which are a type of trichome, and cotton seeds are entirely enclosed and protected in stiff cotton bolls throughout their development (Basra and Malik, 1983). Like legume embryos, which are enclosed in cutinized seed coats and develop in a hypoxic environment (Rolletschek et al., 2002), cotton fibers inside the cotton bolls might also develop in oxygen-deficient conditions. To survive in this hypoxic environment, cotton fiber cells might need to perform anaerobic glycolysis for energy, just as many other plant cells do under similar conditions (Good and Crosby, 1989; Rivoal and Hanson, 1994). It was noted that aerobic glycolysis could protect rapidly proliferating animal cells against oxidative stress by providing more ROS scavengers (Brand, 1997). Combined with the knowledge that H_2O_2 could promote secondary wall deposition in cotton fibers (Potikha et al., 1999), anaerobic glycolysis that is passively adopted by elongating cotton fiber cells might also provide some advantage in the elongation process by reducing ROS generation, impeding secondary wall differentiation.

5. Conclusion

In conclusion, a total of 163 genes encoding different isoforms of 20 enzymes that catalyze glycolytic and fermentation reactions were successfully identified in the diploid Asian cotton (*G. arboreum*) genome. Among them, 125 genes were expressed in Asian cotton plants, and 51 were further identified as expressed in elongating fibers. qRT-PCR confirmed that 6 fiber-specific genes were preferentially accumulated in elongating cotton fibers. Promoter analysis showed that the expression of a majority of these genes was regulated by different phytohormones and by hypoxia. This information not only revealed the systematic transcriptional regulation of glycolysis in cotton fiber cells, but also provided a valuable basis for further functional characterization of glycolytic genes in cotton plants.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.plgene.2015.09.001.

Acknowledgments

This work was sponsored by State Key Laboratory of Cotton Biology Open Fund (CB2015A01) and China Postdoctoral Science Foundation (2014M550074).

References

Bar-Even, A., Flamholz, A., Noor, E., Milo, R., 2012. Rethinking glycolysis: on the biochemical logic of metabolic pathways. Nat. Chem. Biol. 8, 509–517.
Basra, A.S., Malik, C.P., 1983. Dark metabolism of CO_2 during fibre elongation of two cottons differing in fibre lengths. J. Exp. Bot. 34, 1–9.
Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. 57, 289–300.
Brand, K., 1997. Aerobic glycolysis by proliferating cells: protection against oxidative stress at the expense of energy yield. J. Bioenerg. Biomembr. 29, 355–364.
Christofk, H.R., Vander Heiden, M.G., Harris, M.H., Ramanathan, A., Gerszten, R.E., Wei, R., Fleming, M.D., Schreiber, S.L., Cantley, L.C., 2008. The M2 splice isoform of pyruvate kinase is important for cancer metabolism and tumour growth. Nature 452, 230–233.
Claeyssen, E., Rivoal, J., 2007. Isozymes of plant hexokinase: occurrence, properties and functions. Phytochemistry 68, 709–731.
Dennis, D.T., Miernyk, J.A., 1982. Compartmentation of nonphotosynthetic carbohydrate metabolism. Annu. Rev. Plant Physiol. 33, 27–50.
Dhindsa, R.S., Beasley, C.A., Ting, I.P., 1975. Osmoregulation in cotton fiber: accumulation of potassium and malate during growth. Plant Physiol. 56, 394–398.
Emanuelsson, O., Nielsen, H., Brunak, S., von Heijne, G., 2000. Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. J. Mol. Biol. 300, 1005–1016.
Fernie, A.R., Carrari, F., Sweetlove, L.J., 2004. Respiratory metabolism: glycolysis, the TCA cycle and mitochondrial electron transport. Curr. Opin. Plant Biol. 7, 254–261.
Fernie, A.R., Tauberger, E., Lytovchenko, A., Roessner, U., Willmitzer, L., Trethewey, R.N., 2007. Antisense repression of cytosolic phosphoglucomutase in potato (*Solanum tuberosum*) results in severe growth retardation, reduction in tuber number and altered carbon metabolism. Planta 214, 510–520.
Gatenby, R.A., Gillies, R.J., 2004. Why do cancers have high aerobic glycolysis? Nat. Rev. Cancer 4, 4891–4899.
Good, A.G., Crosby, W.L., 1989. Induction of alcohol dehydrogenase and lactate dehydrogenase in hypoxically induced barley. Plant Physiol. 90, 860–866.
Gou, J.Y., Wang, L.J., Chen, S.P., Hu, W.L., Chen, X.Y., 2007. Gene expression and metabolite profiles of cotton fiber during cell elongation and secondary cell wall synthesis. Cell Res. 17, 422–434.
Guan, X., Nah, G., Song, Q., Udall, J.A., Stelly, D.M., Chen, Z.J., 2014. Transcriptome analysis of extant cotton progenitors revealed tetraploidization and identified genome-specific single nucleotide polymorphism in diploid and allotetraploid cotton. BMC Res. Notes 7, 493.
Hovav, R., Udall, J.A., Chaudhary, B., Rapp, R., Flagel, L., Wendel, J.F., 2008. Partitioned expression of duplicated genes during development and evolution of a single cell in a polyploid plant. Proc. Natl. Acad. Sci. U. S. A. 105, 6191–6195.
Klok, E.J., Wilson, I.W., Wilson, D., Chapman, S.C., Ewing, R.M., Somerville, S.C., Peacock, W.J., Dolferus, R., Dennis, E.S., 2002. Expression profile analysis of the low-oxygen response in Arabidopsis root cultures. Plant Cell 14, 2481–2494.
Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J., Higgins, D.G., 2007. Clustal W and clustal X version 2.0. Bioinformatics 23, 2947–2948.
Lescot, M., Déhais, P., Thijs, G., Marchal, K., Moreau, Y., Van de Peer, Y., Rouzé, P., Rombauts, S., 2002. PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. Nucleic Acids Res. 30, 325–327.
Li, F., Fan, G., Wang, K., Sun, F., Yuan, Y., Song, G., Li, Q., Ma, Z., Lu, C., Zou, C., Chen, W., Liang, X., Shang, H., Liu, W., Shi, C., Xiao, G., Gou, C., Ye, W., Xu, X., Zhang, X., Wei, H., Li, Z., Zhang, G., Wang, J., Liu, K., et al., 2014. Genome sequence of the cultivated cotton *Gossypium arboreum*. Nat. Genet. 46, 567–572.
Li, X.R., Wang, L., Ruan, Y.L., 2010. Developmental and molecular physiological evidence for the role of phosphoenolpyruvate carboxylase in rapid cotton fibre elongation. J. Exp. Bot. 61, 287–295.
Liu, F., Vantoai, T., Moy, L.P., Bock, G., Linford, L.D., Quackenbush, J., 2005. Global transcription profiling reveals comprehensive insights into hypoxic response in *Arabidopsis*. Plant Physiol. 137, 1115–1129.
Livak, K.J., Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta Ct}$ method. Methods 25, 402–408.
Lunt, S.Y., Vander Heiden, M.G., 2011. Aerobic glycolysis: meeting the metabolic requirements of cell proliferation. Annu. Rev. Cell Dev. Biol. 27, 441–464.
Mansoor, S., Paterson, A.H., 2012. Genomes for jeans: cotton genomics for engineering superior fiber. Trends Biotechnol. 30, 521–527.
Miernyk, J.A., Dennis, D.T., 1982. Isozymes of the glycolytic enzymes in endosperm from developing castor oil seeds. Plant Physiol. 69, 825–828.
Miller, S.S., Driscoll, B.T., Gregerson, R.G., Gantt, J.S., Vance, C.P., 1998. Alfalfa malate dehydrogenase (MDH): molecular cloning and characterization of five different forms reveals a unique nodule-enhanced MDH. Plant J. 15, 173–184.
Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L., Wold, B., 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat. Methods 5, 621–628.
Mustroph, A., Stock, J., Hess, N., Aldous, S., Dreilich, A., Grimm, B., 2013. Characterization of the phosphofructokinase gene family in rice and its expression under oxygen deficiency stress. Front. Plant Sci. 4, 125.

Naoumkina, M., Hinchliffe, D.J., Turley, R.B., Bland, J.M., Fang, D.D., 2013. Integrated metabolomics and genomics analysis provides new insights into the fiber elongation process in Ligon lintless-2 mutant cotton (*Gossypium hirsutum* L.). BMC Genomics 14, 155.

Plaxton, W.C., 1996. The organization and regulation of plant glycolysis. Annu. Rev. Plant Physiol. Plant Mol. Biol. 47, 185–214.

Potikha, T.S., Collins, C.C., Johnson, D.I., Delmer, D.P., Levine, A., 1999. The involvement of hydrogen peroxide in the differentiation of secondary walls in cotton fibers. Plant Physiol. 119, 849–858.

Pu, L., Li, Q., Fan, X., Yang, W., Xue, Y., 2008. The R2R3 MYB transcription factor GhMYB109 is required for cotton fiber development. Genetics 180, 811–820.

Puranik, S., Sahu, P.P., Srivastava, P.S., Prasad, M., 2012. NAC proteins: regulation and role in stress tolerance. Trends Plant Sci. 17, 369–381.

Rivoal, J., Hanson, A.D., 1994. Metabolic control of anaerobic glycolysis (overexpression of lactate dehydrogenase in transgenic tomato roots supports the Davies-Roberts hypothesis and points to a critical role for lactate secretion. Plant Physiol. 106, 1179–1185.

Rolletschek, H., Borisjuk, L., Koschorreck, M., Wobus, U., Weber, H., 2002. Legume embryos develop in a hypoxic environment. J. Exp. Bot. 53, 1099–1107.

Ruan, Y.L., Llewellyn, D.J., Furbank, R.T., 2003. Suppression of sucrose synthase gene expression represses cotton fiber cell initiation, elongation, and seed development. Plant Cell 15, 952–964.

Shi, Y.H., Zhu, S.W., Mao, X.Z., Feng, J.X., Qin, Y.M., Zhang, L., Cheng, J., Wei, L.P., Wang, Z.Y., Zhu, Y.X., 2006. Transcriptome profiling, molecular biological, and physiological studies reveal a major role for ethylene in cotton fiber cell elongation. Plant Cell 18, 651–664.

Sturm, A., 1999. Invertases. Primary structures, functions, and roles in plant development and sucrose partitioning. Plant Physiol. 121, 1–8.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S., 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. 28, 2731–2739.

Todeschini, A.L., Georges, A., Veitia, R.A., 2014. Transcription factors: specific DNA binding and specific gene regulation. Trends Genet. 30, 211–219.

Usadel, B., Obayashi, T., Mutwil, M., Giorgi, F.M., Bassel, G.W., Tanimoto, M., Chow, A., Steinhauser, D., Persson, S., Provart, N.J., 2009. Co-expression tools for plant biology: opportunities for hypothesis generation and caveats. Plant Cell Environ. 32, 1633–1651.

Van der Straeten, D., Rodrigues-Pousada, R.A., Goodman, H.M., Van Montagu, M., 1991. Plant enolase: gene structure, expression, and evolution. Plant Cell 3, 719–735.

Walker, J.C., Howard, E.A., Dennis, E.S., Peacock, W.J., 1987. DNA sequences required for anaerobic expression of the maize alcohol dehydrogenase 1 gene. Proc. Natl. Acad. Sci. U. S. A. 84, 6624–6628.

Wang, L., Li, X.R., Lian, H., Ni, D.A., He, Y.K., Chen, X.Y., Ruan, Y.L., 2010. Evidence that high activity of vacuolar invertase is required for cotton fiber and *Arabidopsis* root elongation through osmotic dependent and independent pathways, respectively. Plant Physiol. 154, 744–756.

Xiao, Y.H., Li, D.M., Yin, M.H., Li, X.B., Zhang, M., Wang, Y.J., Dong, J., Zhao, J., Luo, M., Luo, X.Y., Hou, L., Hu, L., Pei, Y., 2010. Gibberellin 20-oxidase promotes initiation and elongation of cotton fibers by regulating gibberellin synthesis. J. Plant Physiol. 167, 829–837.

Yazawa, T., Kawahigashi, H., Matsumoto, T., Mizuno, H., 2013. Simultaneous transcriptome analysis of sorghum and *Bipolaris sorghicola* by using RNA-seq in combination with de novo transcriptome assembly. PLoS One 8, e62460.

Zancan, P., Sola-Penna, M., Furtado, C.M., Da Silva, D., 2010. Differential expression of phosphofructokinase-1 isoforms correlates with the glycolytic efficiency of breast cancer cells. Mol. Genet. Metab. 100, 372–378.

Zhang, B., Yang, Y.W., Zhang, Y., Liu, J.Y., 2013. A high-confidence reference dataset of differentially expressed proteins in elongating cotton fiber cells. Proteomics 13, 1159–1163.

Zhang, M., Zheng, X., Song, S., Zeng, Q., Hou, L., Li, D., Zhao, J., Wei, Y., Li, X., Luo, M., Xiao, Y., Luo, X., Zhang, J., Xiang, C., Pei, Y., 2011. Spatiotemporal manipulation of auxin biosynthesis in cotton ovule epidermal cells enhances fiber yield and quality. Nat. Biotechnol. 29, 453–458.

Zou, C., Lu, C., Shang, H., Jing, X., Cheng, H., Zhang, Y., Song, G., 2013. Genome-wide analysis of the *Sus* gene family in cotton. J. Integr. Plant Biol. 55, 643–653.