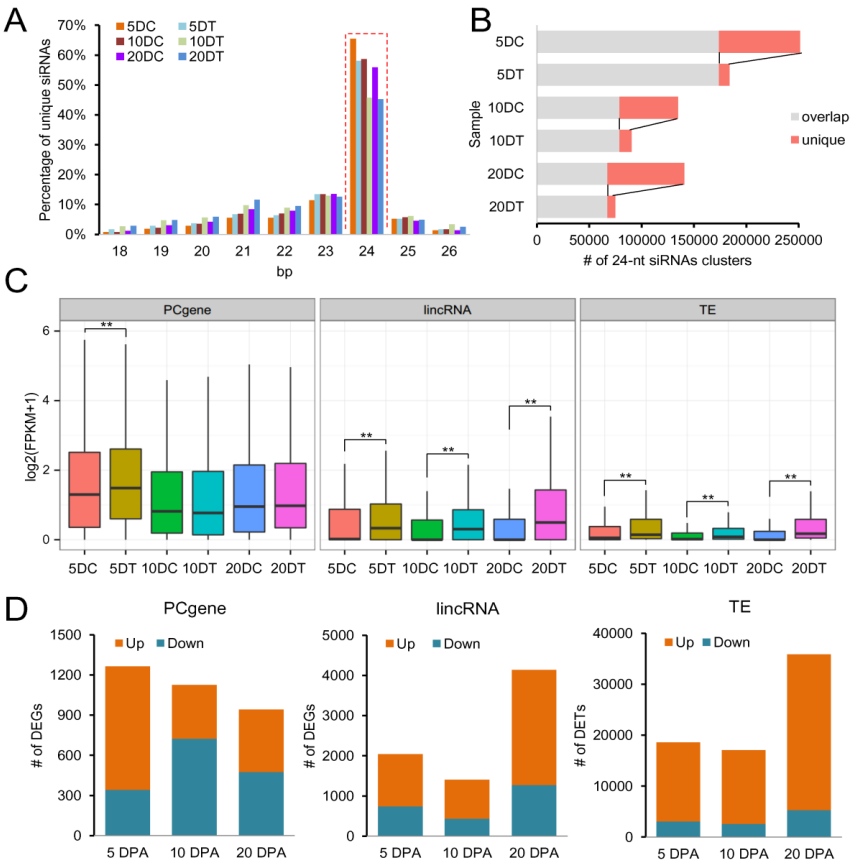
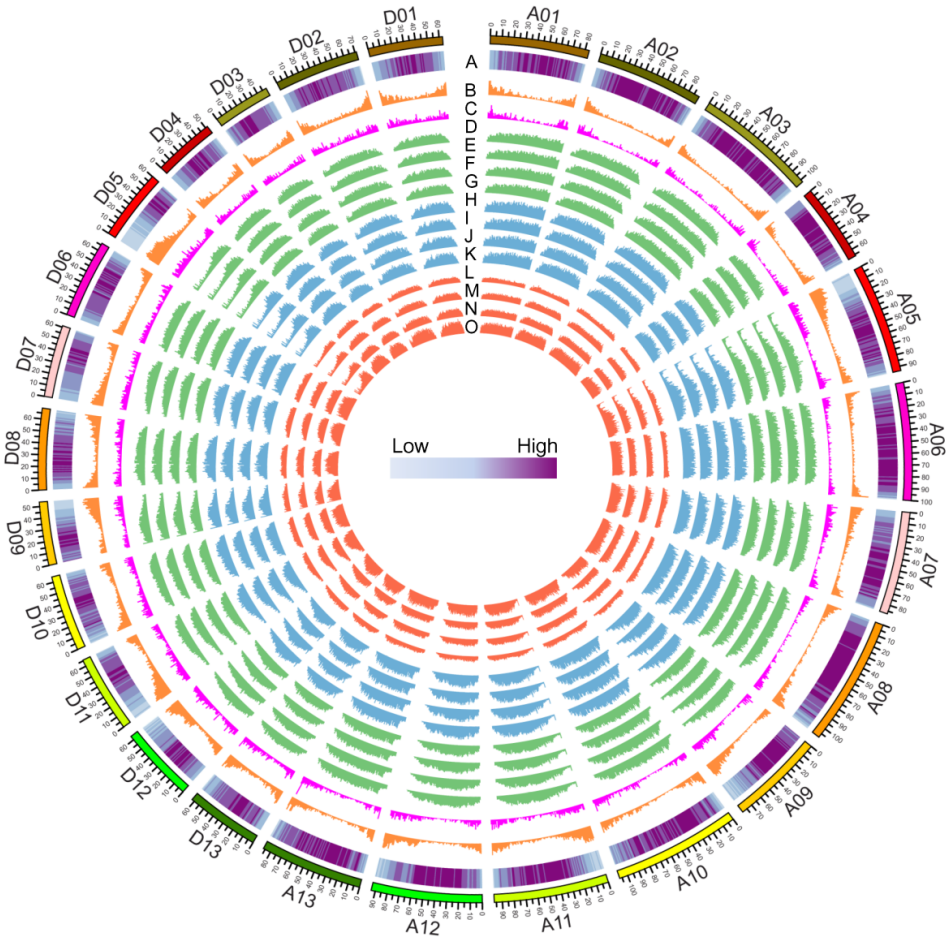
Supplementary Figures



Supplementary Figure S1 Disruption of small RNAs and gene expression in zebularine-treated cotton ovules.

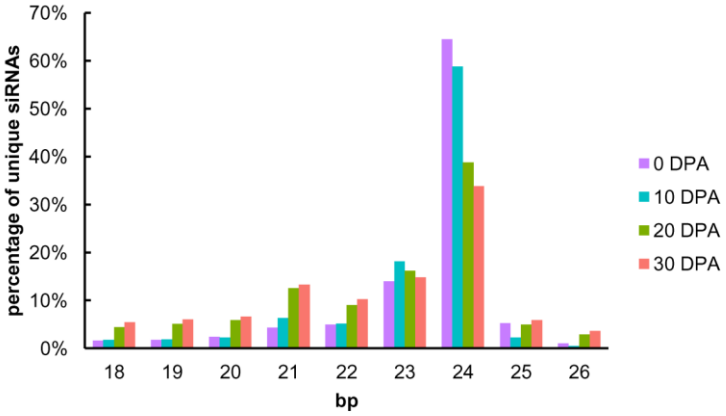
(A) Percentage of unique small RNAs with lengths ranging from 18-26 nt. The 24-nt unique small RNA is highlighted with the red dotted box. (B) Identification of 24-nt siRNA clusters in each sample. (C) Expression levels of protein-coding gene (PCgene), long intergenic non-coding RNA (lincRNA) and transposable element (TE), which are marked by 24-nt siRNAs. The samples include ovules of control and treatment at 5 DPA (5DC and 5DT), fibers of control and treatment at 10 DPA (10DC and 10DT) and fibers of control and treatment at 20 DPA (20DC and 20DT). (D) Identification of differentially expressed PCgenes, lincRNAs, and TEs after zebularine treatment of cotton ovule culture.



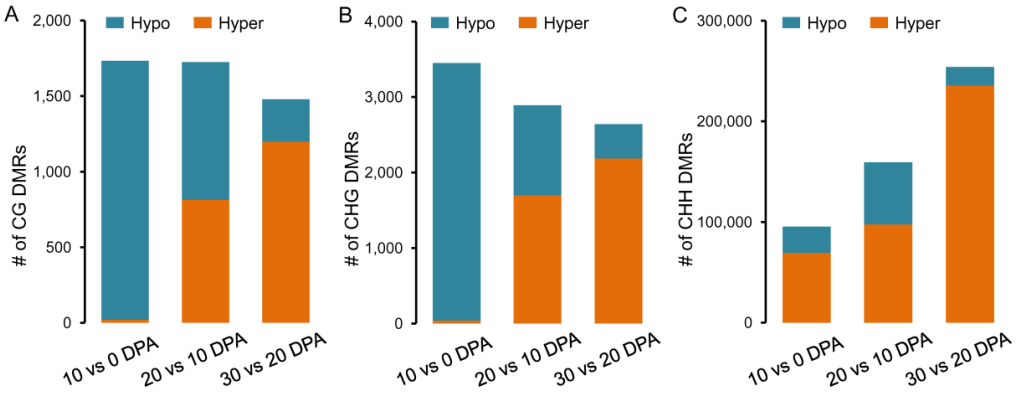
Supplementary Figure S2 Genomic landscape of DNA methylation in cotton fiber development.

The outer track shows the 26 chromosomes in allotetraploid *G. barbadense* genome (A01-A13 of At subgenome and D01-D13 of Dt subgenome). The inner 14 colored histogram tracks show the numbers of genes and methylation levels. (A) Heatmap shows the TE content. (B) The content of protein-coding genes. (C) The content of long intergenic non-coding RNAs. (D-G) CG methylation levels in ovules at 0 DPA, and fibers at 10 DPA, 20 DPA and 30 DPA, respectively. (H-K) CHG methylation levels in ovules at 0 DPA, and fibers at 10 DPA, 20 DPA and 30 DPA, respectively. (L-O) CHH methylation levels in ovules at 0 DPA, and fibers at 10 DPA, 20 DPA and 30 DPA, respectively. The data for each chromosome were counted in 1 Mb windows

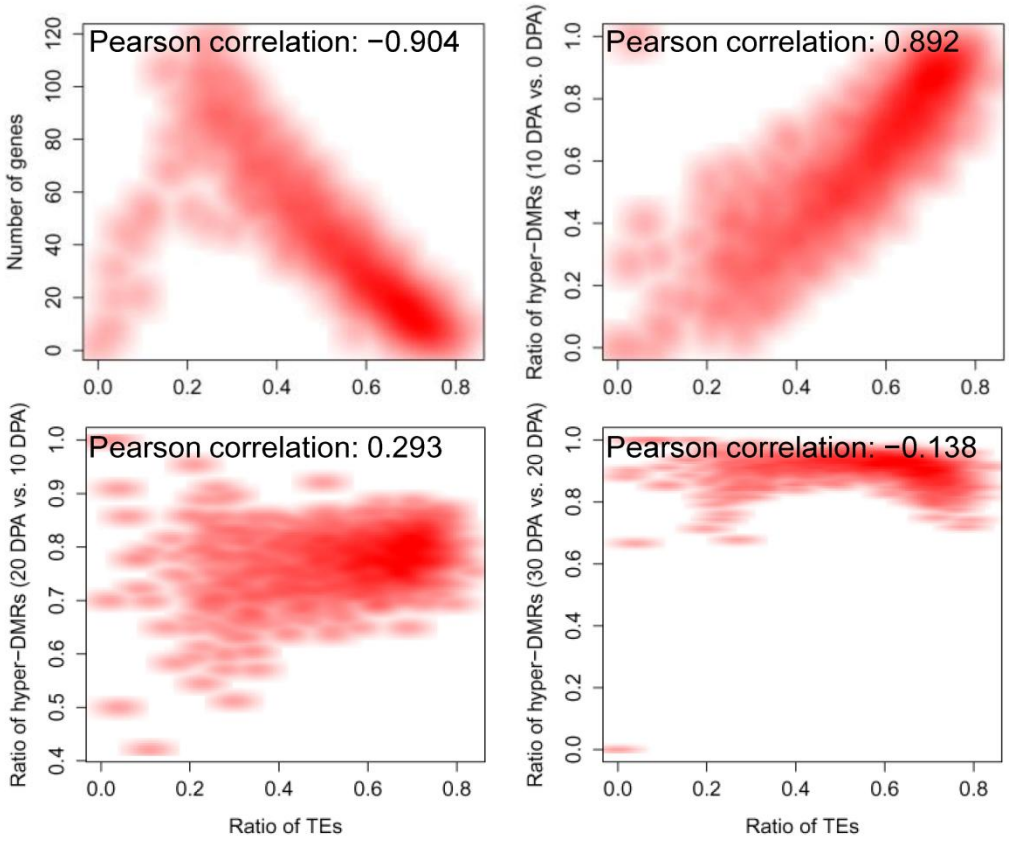
sliding 200 kb. For tracks D-O, the CG, CHG and CHH methylation levels in each windows were determined by the ratio of methylated cytosines to total sequenced cytosines.



Supplementary Figure S3 The proportions of unique small RNAs at different fiber development stages.

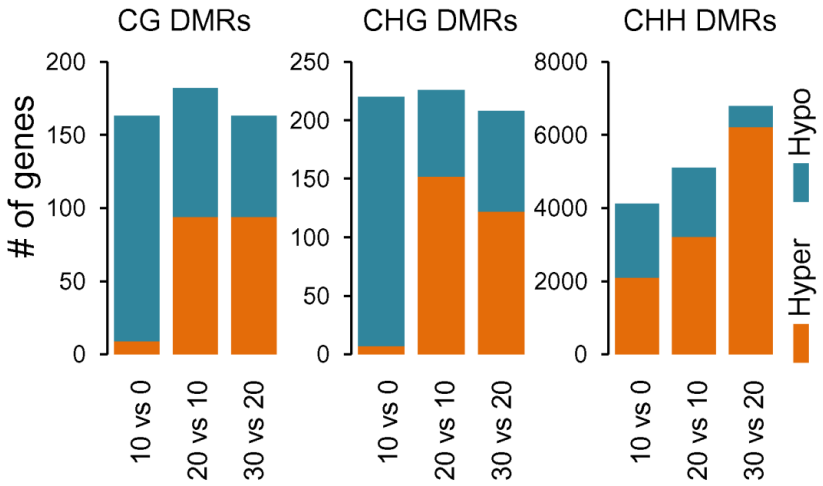


Supplementary Figure S4 The numbers of hypomethylated (Hypo) and hypermethylated (Hyper) context-specific differentially methylated regions (DMRs) between consecutive developmental stages.

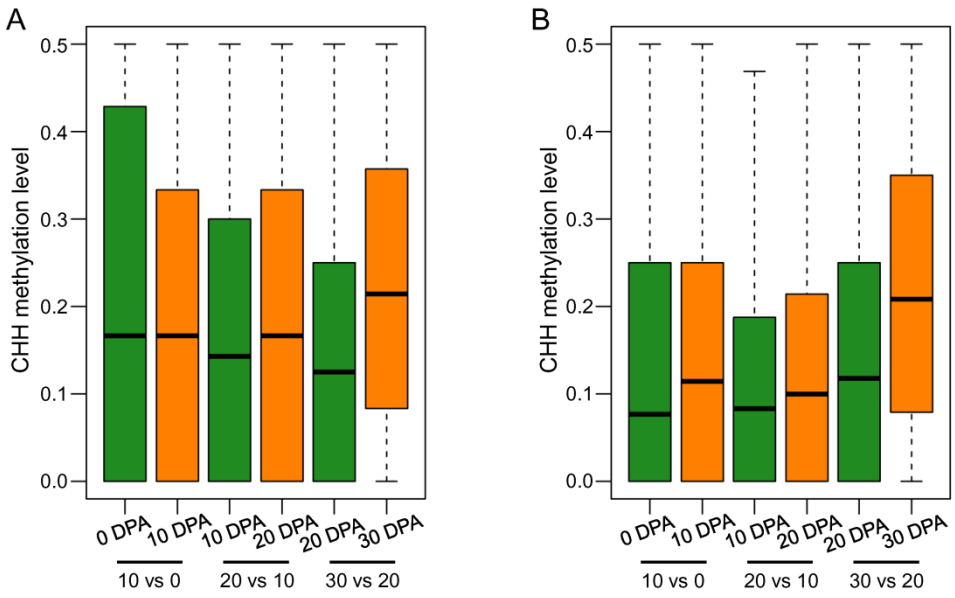


Supplementary Figure S5 Correlations between TE content and the numbers of DMRs in genomic 1 Mb regions.

The *G. barbadense* genome was divided into non-overlapping 1 Mb regions. For each region, the total numbers of CG, CHG and CHH DMRs were counted. The Pearson correlation method was applied to calculate the correlation index.

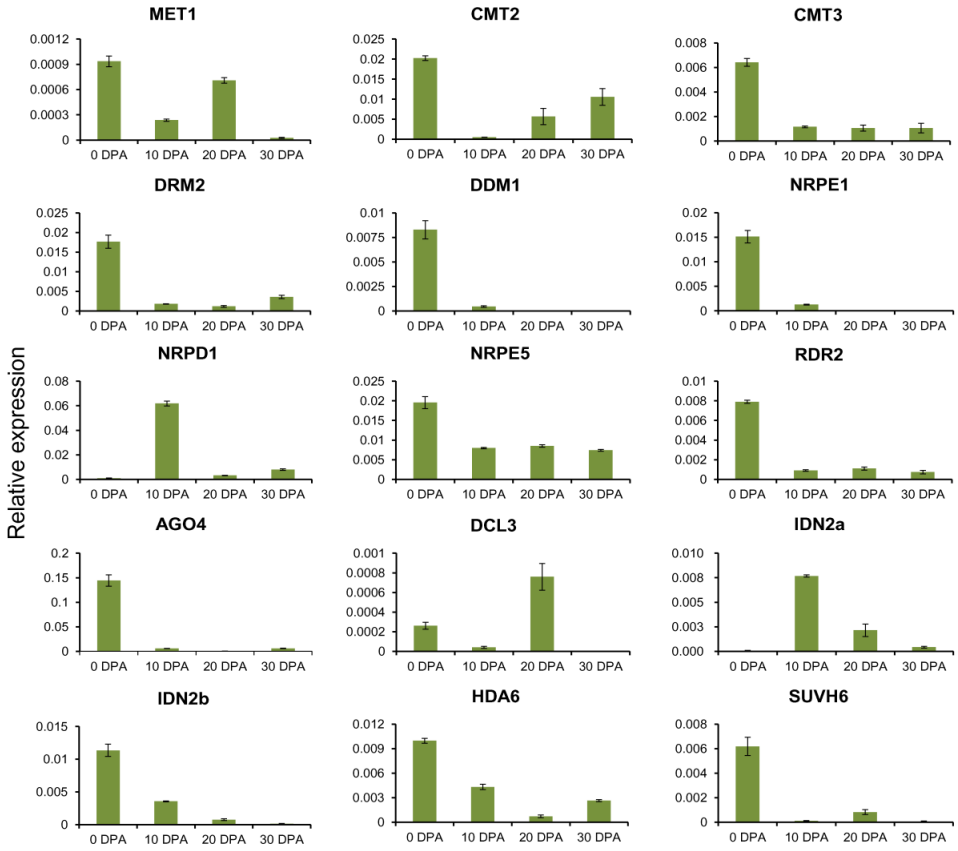


Supplementary Figure S6 Number of DMRs-marked protein-coding genes between two consecutive stages.



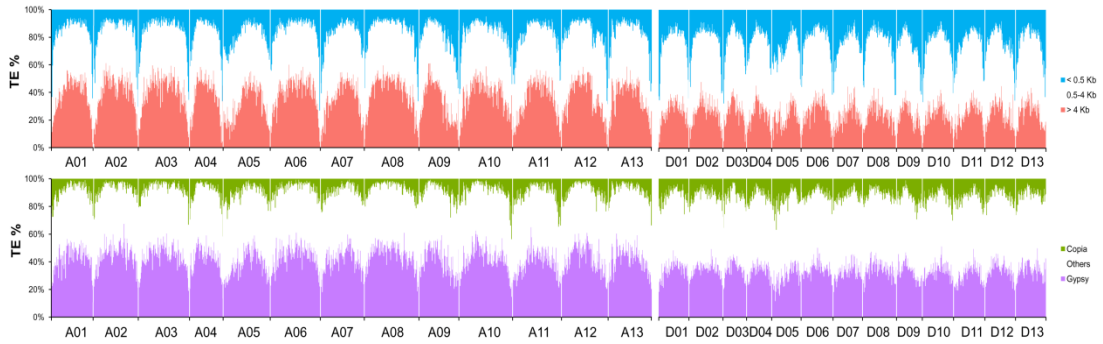
Supplementary Figure S7 The CHH methylation levels in 24-nt siRNA production regions.

The *G. barbadense* genome was divided into non-overlapping 100-bp regions. For each region, the 24-nt siRNA level and CHH methylation level were calculated. (A) The CHH methylation levels in down-regulated siRNA production regions. (B) The CHH methylation levels in up-regulated siRNA production regions.



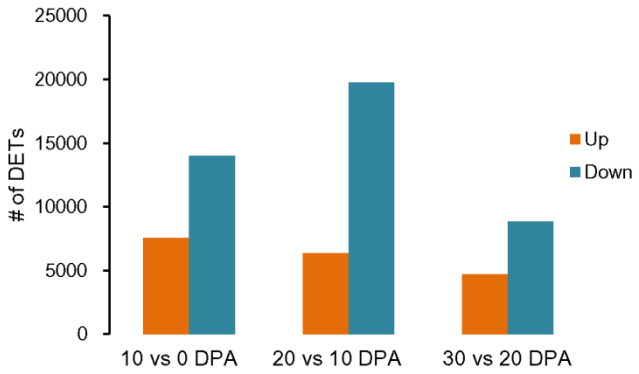
Supplementary Figure S8 Expression analysis of genes including DNA methytransferases and some key genes in RNA-directed DNA methylation pathway using real-time PCR.

These genes include Methyltransferase 1 (*MET1*), Chromomethylase 2 (*CMT2*), Chromomethylase 3 (*CMT3*), Domains Rearranged Methyltransferase 2 (*DRM2*), Chromatin Remodeling 1 (*DDM1*), Nuclear RNA polymerase D1B (*NRPE1*), Nuclear RNA polymerase D1A (*NRPD1*), RPB5-like subunits (*NRPE5*), RNA-dependent RNA polymerase 2 (*RDR2*), Argonaute 4 (*AGO4*), Dicer-like 3 (*DCL3*), Involved in de novo (*IDN2*), Histone Deacetylase 6 (*HDA6*) and *SUVH6*. Gene names are based on the definitions in *Arabidopsis*. The expression levels were normalized to UB7. Error bars, ± standard deviation (SD) of three biological replicates.



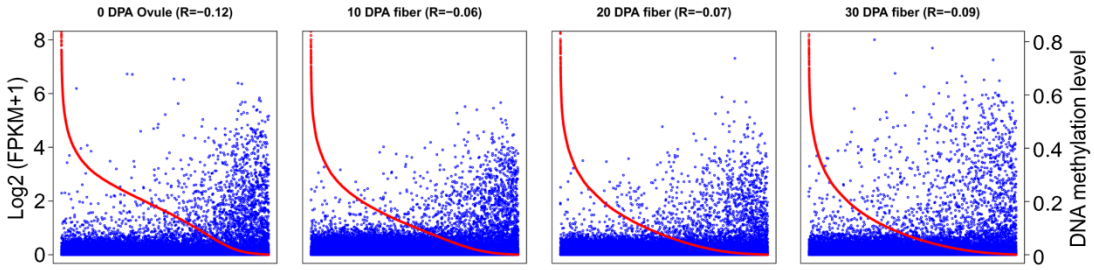
Supplementary Figure S9 Genome-wide TEs in the At (A01-A13) and Dt (D01-D13) subgenomes of *G. barbadense* genome.

The upper histogram shows the ratios of TEs with different lengths (> 4 kb, long TEs in red; 0.5-4 kb, middle-sized TEs in white; < 0.5 kb, short TEs in blue). The lower histogram shows the ratios of *Gypsy* and *Copia* long TEs (*Cypsy* in purple, *Copia* in grassgreen, other long TEs in white).



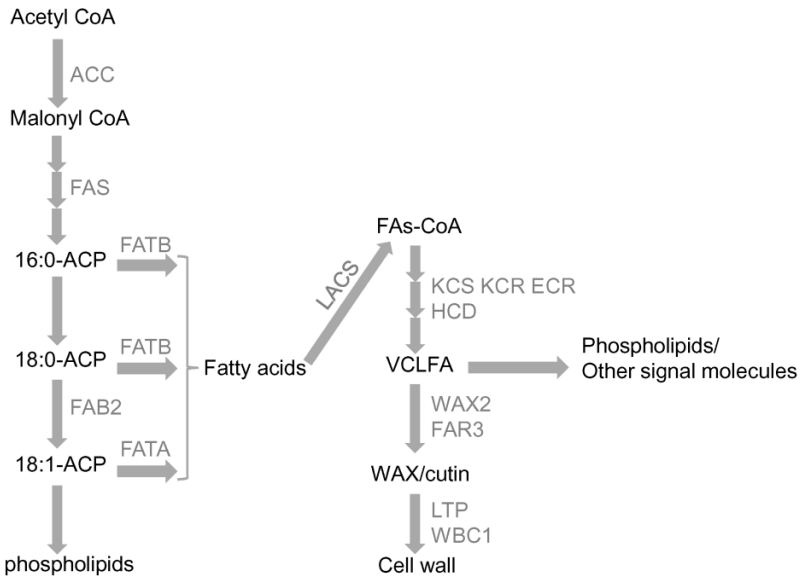
Supplementary Figure S10 The numbers of differentially expressed long TEs in two consecutive developmental stages.

The differentially expressed long TEs were identified using the DESeq package with expression changes of at least two-fold (FDR < 0.01). The up-regulated genes are shown in brown and down-regulated genes are shown in navy blue.



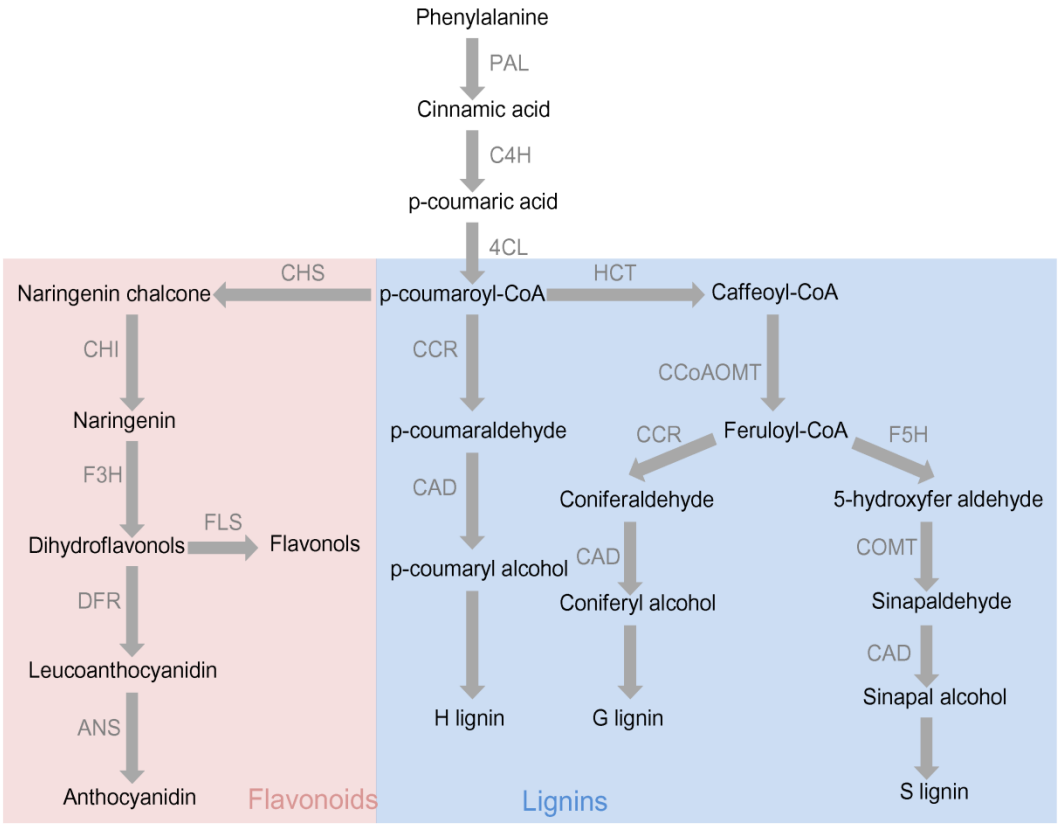
Supplementary Figure S11 Correlations of gene expression and DNA methylation levels in the promoter 2 kb regions.

The red lines show gene expression levels and blue dots show DNA methylation levels. The expression levels of genes were sorted in this analysis.



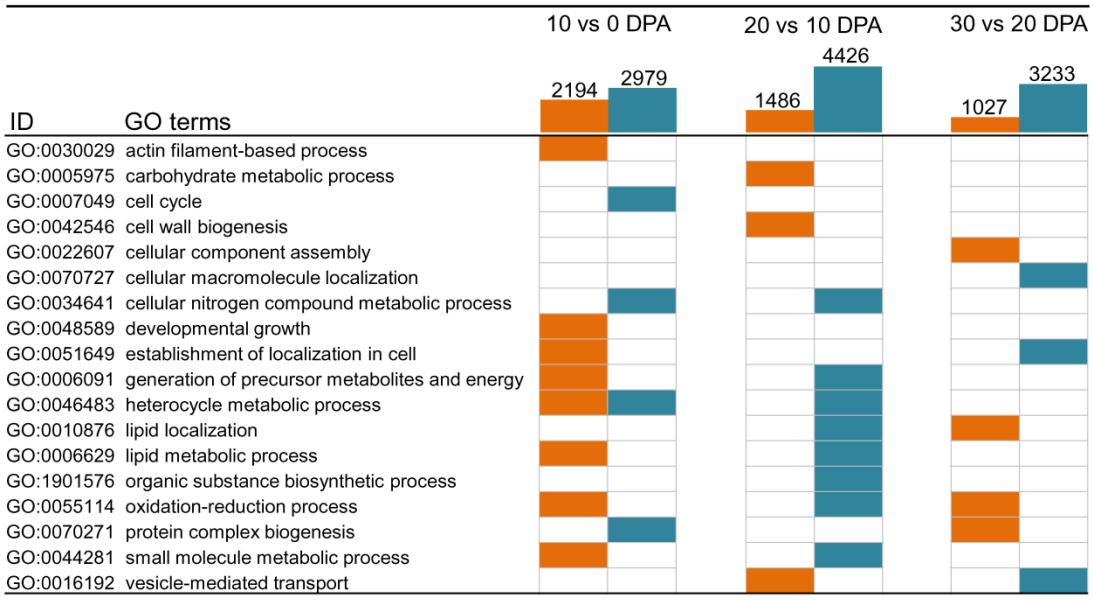
Supplementary Figure S12 The lipid metabolic pathway.

The key genes are indicated for each metabolic process. Genes include Acetyl-CoA Carboxylase (*ACC*), Fatty Acid Synthase (*FAS*), Fatty Acyl-ACP Thioesterases A/B (*FATA/B*), Long Chain Acyl-CoA Synthetase (*LACS*), 3-ketoacyl-CoA Synthase (*KCS*), Beta-ketoacyl Reductase (*KCR*), Enoyl-CoA Reductase (*ECR*), 3-hydroxyacyl-CoA Dehydratase (*HCD*), *WAX2*, Jojoba acyl CoA reductase-related male sterility protein 3 (*FAR3*), Lipid Transfer Protein 3 (*LTP3*) and ATP-binding Cassette Transporter (*WBC1*).



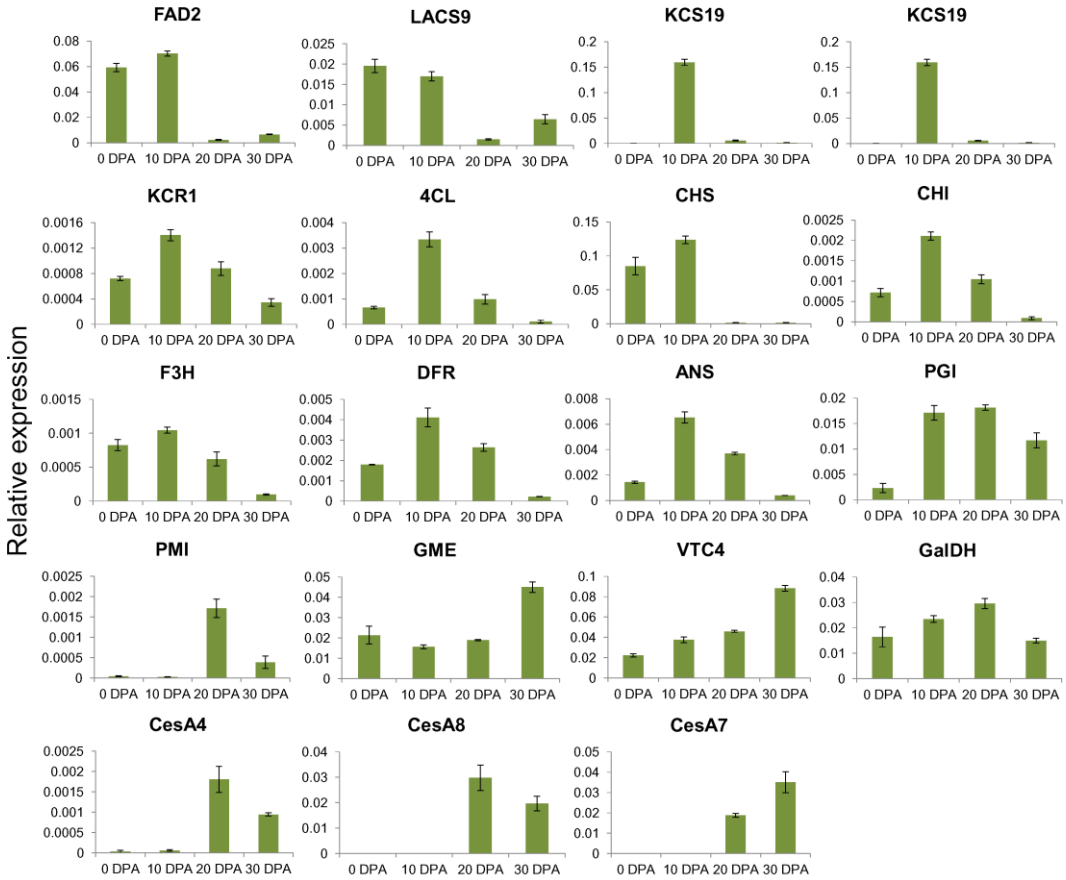
Supplementary Figure S13 The phenylpropanoid metabolic pathway.

The key genes are indicated for each metabolic process. Genes include Phenylalanine Ammonia-Lyase (*PAL*), Cinnamate-4-Hydroxylase (*C4H*), 4-coumarate:CoA Ligase (*4CL*), Chalcone Synthase (*CHS*), Chalcone Isomerase (*CHI*), Flavanone 3-hydroxylase (*F3H*), Flavonol Synthase (*FLS*), Dihydroflavonol-4-reductase (*DFR*), Leucoanthocyanidin Dioxygenase (*ANS*), Hydroxycinnamoyltransferase (*HCT*), Cinnamoyl CoA Reductase (*CCR*), Cinnamyl Alcohol Dehydrogenase (*CAD*), O-methyltransferase (*COMT*), Caffeoyl-CoA O-methyltransferase (*CCoAOMT*).



Supplementary Figure S14 A Gene Ontology (GO) analysis of differentially expressed genes in cotton fiber development.

The upper track shows the number of differentially expressed genes. For each comparison, the up-regulated genes are indicated in brown and down-regulated genes are indicated in blue. The down table shows the significantly enriched GO terms in each color (up-regulated genes in brown and down-regulated genes in blue) (FDR < 0.01).



Supplementary Figure S15 Expression analysis of key genes using real-time PCR. The full descriptions of these genes are found in Supplementary Table S6. The expression levels were normalized to UB7. Error bars, ± standard deviation (SD) of three biological replicates.

Supplementary Tables

Supplementary Table S1 Summary of raw and clean RNA-Seq data in ovule-cultures.

Sample	# of raw read pairs	# of clean pair-end reads	# of clean single-end reads
5DPA control (5DC)	22,363,150	19,589,279	2,273,918
5DPA treat (5DT)	24,126,172	20,899,655	2,571,478
10DPA control (10DC)	27,346,414	24,101,947	2,598,232
10DPA treat (10DT)	33,447,273	28,091,291	4,273,333
20DPA control (20DC)	14,799,197	12,746,767	1,493,975
20DPA treat (20DT)	22,235,350	19,044,364	2,530,510

Supplementary Table S2 Summary of small RNA-Seq data in ovule-cultures.

Sample	Total unique reads	Total raw reads
5DPA control (5DC)	3,896,140	11,745,767
5DPA treat (5DT)	3,036,640	11,584,519
10DPA control (10DC)	2,977,450	11,585,674
10DPA treat (10DT)	2,305,323	11,410,923
20DPA control (20DC)	2,860,362	11,642,482
20DPA treat (20DT)	1,952,624	11,588,692

Supplementary Table S3 Summary of bisulfite-converted DNA sequencing data in *G. barbadense*.

Sample	# of read pairs	Read length	Genome coverage	Unique mapping	Multiple mapping
Ovule 0 DPA	269,428,371	101 bp	21.5	62.6%	27.9%
Fiber 10 DPA	253,185,627	101 bp	20.2	51.6%	36.1%
Fiber 20 DPA	248,256,276	101 bp	19.8	49.7%	20.5%
Fiber 30 DPA	243,307,801	101 bp	19.4	53.5%	26.4%

Supplementary Table S4 Summary of raw and clean MNase-Seq data.

Sample	# of raw pair-end reads	# of clean pair-end reads	# of clean single-end reads
0 DPA	53,810,732	39,868,117	9,427,964
10 DPA	48,489,717	31,302,062	12,148,925
20 DPA	50,447,156	34,584,898	10,491,238
30 DPA	39,576,576	24,255,866	9,357,890

Supplementary Table S5 The number of H3K9me2 ChIP peaks.

Sample	# of ChIP peaks
0 DPA	10,102
10 DPA	13,518
20 DPA	14,256
30 DPA	14,774

Supplementary Table S6 Summary of the full names of key genes.

Name	Description
CWIN	cell wall invertase
SUS	sucrose synthase
HXK	hexokinase
PGI	phosphoglucose isomerase
PMI	mannose-6-phosphate isomerase
PMM	phosphomannomutase
VTC	mannose-1-phosphate guanylyltransferase
GME	GDP-D-mannose 3',5'-epimerase
GalDH	2,5-diketo-D-gluconic acid reductase
GLDH	L-galactono-1,4-lactone dehydrogenase
FAD	fatty acid desaturase
LACS	long chain acyl-CoA synthetase
KCS	3-ketoacyl-CoA synthase
KCR	beta-ketoacyl reductase
GSH	glutamate-cysteine ligase
UGDH	UDP-glucose-6-dehydrogenase
UGD	UDP-glucuronate decarboxylase

GAE	UDP-D-glucuronate-4-epimerase
XTH	xyloglucan endotransglycosylase
4CL	4-coumarate:CoA ligase
CHS	chalcone synthase
CHI	chalcone isomerase
F3H	flavanone 3-hydroxylase
DFR	dihydroflavonol-4-reductase
ANS	leucoanthocyanidin dioxygenase
CesA	cellulose synthase

Supplementary Table S7 Summary of the levels of metabolites during cotton fiber elongation and transition to secondary cell wall synthesis in *G. barbadense*.

Pathway	Name	10 DPA	15 DPA	18 DPA	21 DPA	28 DPA
Lipids, Free fatty acid	linolenate	1.94	3.65	1.52	1.42	0.31
Lipids, Glycerolipids	1-oleoylglycerol	1.10	1.85	1.43	0.87	0.19
Lipids, Oxylipins	9,10-epoxystearate	1.02	1.91	0.82	0.92	0.28
Lipids, Phospholipids	glycerol 3-phosphate	1.20	1.31	0.84	0.91	1.25
Lipids, Sphingolipid	phytosphingosine	1.96	1.49	0.93	0.47	0.12
Lipids, Sterols	squalene	1.52	1.17	0.69	0.59	0.96
Flavonoids	catechin	2.56	1.94	1.27	0.94	1.05
Flavonoids	procyanidin B1	0.41	1.18	0.70	0.40	0.31
Flavonoids	dihydrokaempferol	0.51	0.41	0.38	0.38	0.38
Flavonoids	naringenin-7-O-glucoside	1.56	0.81	0.35	0.40	0.35
Flavonoids	leucocyanidin	2.65	1.79	1.05	0.72	0.82
Ascorbate metabolism	glucarate	0.59	1.37	1.11	1.09	3.07
Ascorbate metabolism	ascorbate	0.91	0.83	4.15	5.20	6.81
Ascorbate metabolism	threonate	0.46	1.04	0.59	0.67	1.28
Ascorbate metabolism	gulono-1,4-lactone	1.05	1.13	0.56	0.58	1.38