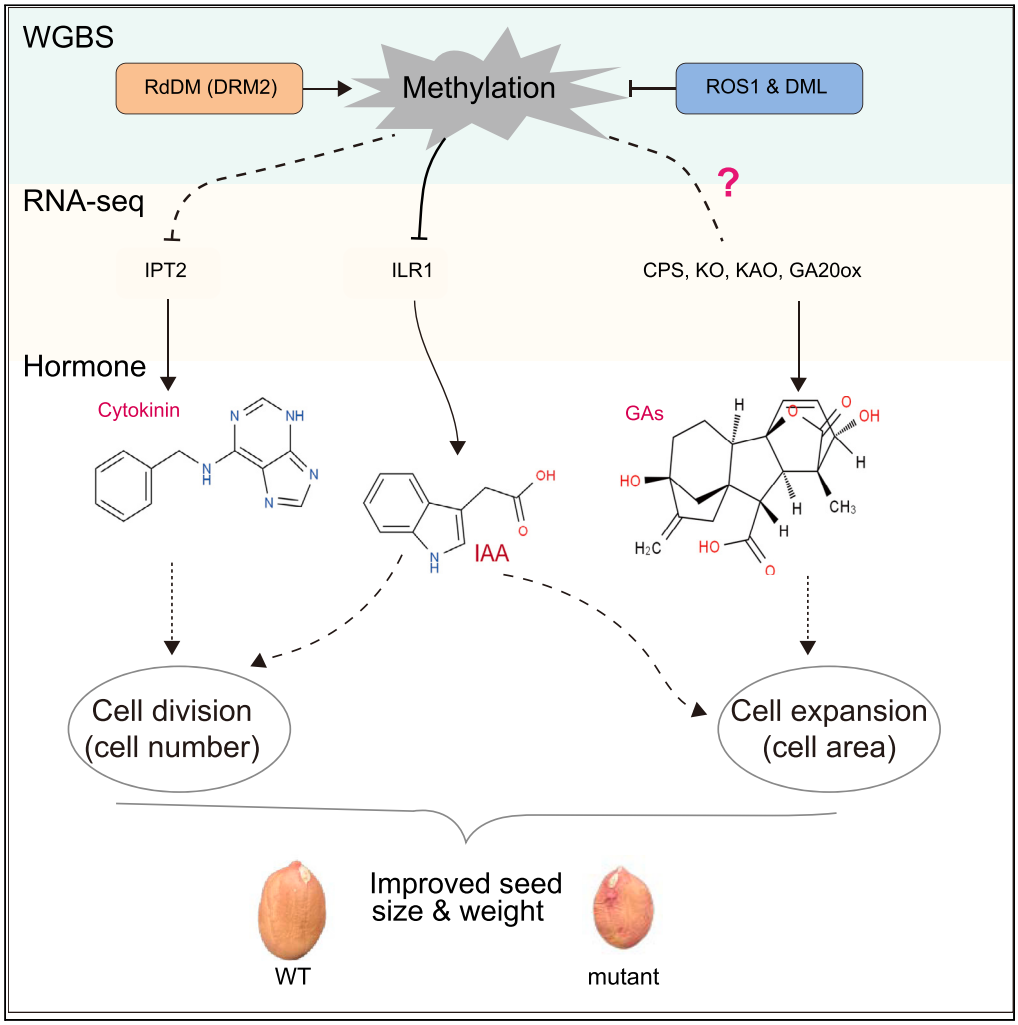
iScience

CellPress
OPEN ACCESS

Article

Dynamic DNA methylation modification in peanut seed development



Zhongfeng Li, Qian Liu, Kai Zhao, ..., Ding Qiu, Yu Zhao, Dongmei Yin

zhaoyu@mail.hzau.edu.cn (Y.Z.)
yindm@henau.edu.cn (D.Y.)

Highlights

The DNA methylation profiles for developmental peanut seeds at two stages

Globally elevated methylation and enhanced CHH methylation during seed development

Identification of differentially methylated regions associated with seed development

Negative association of methylation in promoter or gene body with expression levels

Li et al., iScience 26, 107062
July 21, 2023
https://doi.org/10.1016/j.isci.2023.107062

iScience

CellPress
OPEN ACCESS

Article

Dynamic DNA methylation modification in peanut seed development

Zhongfeng Li,[1,3] Qian Liu,[2,3] Kai Zhao,[1,3] Di Cao,[1] Zenghui Cao,[1] Kunkun Zhao,[1] Qian Ma,[1] Gaidan Zhai,[1] Sasa Hu,[1] Zhan Li,[1] Kuopeng Wang,[1] Fangping Gong,[1] Xingli Ma,[1] Xingguo Zhang,[1] Rui Ren,[1] Ding Qiu,[1] Yu Zhao,[2,*] and Dongmei Yin[1,4,*]

SUMMARY

Cytosine methylation is an important epigenetic modification involved in regulation of plant development. However, the epigenetic mechanisms governing peanut seed development remain unclear. Herein, we generated DNA methylation profiles of developmental seeds of peanut H2014 and its smaller seed mutant H1314 at 15 and 60 days after pegging (DAP, S1, S4). Accompanying seed development, globally elevated methylation was observed in both lines. The mutant had a higher methylation level of 31.1% than wild type at S4, and 27.1–35.9% of the differentially methylated regions (DMRs) between the two lines were distributed in promoter or genic regions at both stages. Integrated methylome and transcriptome analysis revealed important methylation variations closely associated with seed development. Furthermore, some genes showed significantly negative correlation of expression with the methylation level within promoter or gene body. The results provide insights into the roles of DNA methylation in peanut seed development.

INTRODUCTION

Cultivated peanut (*Arachis hypogaea* L.), an important cash crop widely grown in ~110 countries/regions, with a total annual area harvested of 32.70 million hectares (FAOSTAT, 2021), greatly increases income of local farmers. Peanut yield-related agronomic traits including seed size/weight usually display continuous phenotypic variation, and are controlled by multiple quantitative trait loci (QTLs).[1–3] Numerous genomic regions have been identified which control seed yield in peanut using forward genetics, many of which are associated with seed size/weight phenotypes. However, few causal genes controlling seed size/weight-associated QTLs have been characterized.

Plant hormones have been proven to be involved in regulating seed size/weight in the model plants Arabidopsis, rice and other non-model crops,[4] which can regulate seed size/weight through either maternal or the zygotic tissues. Cytokinins (CKs), one of several types of endogenous phytohormones, play vital roles in various plant development processes such as shoot and root growth, flower and female gametophyte development, and control of cell division and differentiation,[5] and they also greatly influence seed yield.[6] For example, increasing accumulation of CKs in inflorescence meristems, following reduced expression of *OsCKX2* and the wheat ortholog *TaCKX6-D1* (cytokinin oxidase/dehydrogenase), leads to higher grain yield in both rice[6] and hexaploid cultivated wheat.[7] In addition, moderately enhanced CK levels can significantly increase seed size and yield in cotton.[8]

Auxins are also key regulators of plant seed development. TGW6 (THOUSAND GRAIN WEIGHT 6), an enzyme with indole-3-acetic acid (IAA)-glucose hydrolase activity, positively modulates free IAA supply by hydrolyzing IAA-glucose in rice grains, and functional loss of TGW6 enhances rice grain length and weight.[9] Recently, AUXIN RESPONSE FACTOR 4 (OsARF4) has been reported to negatively regulate both grain size and weight through the auxin signaling pathway in rice.[10] However, how phytohormones affect peanut seed development remains elusive.

Cytosine methylation of genomic DNA, adding of a methyl group to the 5′ position of a cytosine, is an important epigenetic modification having multifaceted roles in plants, such as maintenance of genome

[1]College of Agronomy, Henan Agricultural University, Zhengzhou 450046, Henan Province, China

[2]National Key Laboratory of Crop Genetic Improvement, Hubei Hongshan Laboratory, Huazhong Agricultural University, Wuhan 430070, Hubei Province, China

[3]These authors contributed equally

[4]Lead contact

*Correspondence: zhaoyu@mail.hzau.edu.cn (Y.Z.), yindm@henau.edu.cn (D.Y.)

https://doi.org/10.1016/j.isci.2023.107062

stability, modulation of gene expression, developmental regulation, and environmental responses.[11,12] In plants, DNA cytosine-5 methylation occurs predominantly in CG, CHG, and CHH contexts (H represents A, T, or C). Cytosine methylation has been extensively studied in crops and other species. For example, Soybean DNA methylation profilings in roots, stems, leaves, and cotyledons of immature seeds were carried out, and some of differentially methylated regions (DMRs) among organs were correlated with elevated expression of neighboring genes.[13] Xing et al. (2015) characterized the DNA methylation genome in embryo and endosperm (2–9 days after pollination) tissues, and revealed the key roles of DNA methylation during rice seed development.[14] In addition, DNA methylomes of maize embryo and endosperm were also examined,[15] and DNA methylation played critical roles in regulation of wheat preharvest sprouting resistance.[16] Notably, dynamic DNA methylation was found to be critical for proper orange fruit development[17] or postharvest ripening of tomato.[18]

DNA methylation has also been studied in peanut. In the genome assembly of the allotetraploid cultivated peanut *A. hypogaea*, cytosine-5 methylation was analyzed in young leaves, and genic methylation patterns typical of plants were observed with decreased methylation in transcribed regions and at transcription start and end sites.[19] Recently, Zhou X et al. (2021) examined epigenetic signatures associated with peanut allergy using targeted next-generation bisulfite sequencing.[20] However, dynamic DNA methylation in developing seeds has not yet been examined, and whether DNA methylation affects seed development in peanut remains unclear.

Herein, we generated single-base resolution maps of DNA methylation for developmental seeds from peanut accession H2014 and its smaller seed mutant using whole-genome bisulfite sequencing (WGBS), and characterized dynamic methylation during seed development of both peanut lines. Based on an integrated analysis of methylome, transcriptome and endogenous phytohormones, we explored the correlation between gene expression and methylation, and the differentially expressed genes (DEGs) with differential methylation closely associated with seed size and weight regulation in peanut.

RESULTS

Seed phenotypes of wild type (H2014) and mutant (H1314) peanut accessions

The mutant lines were derived from an EMS-mutagenized M_2 population of the cultivated peanut accession H2014.[21] A large variation in seed phenotype was observed between the mutant and its parent. Wild type plants had a pod weight of ~53 g on average at maturity, compared with only ~30 g for the mutant (57% of wild type; Figures 1A and 1B). However, the number of pods per plant was not significantly different between the two accessions (Figure 1B). Fresh seeds from both accessions were collected and further examined at young (15 DAP) and expansion (60 DAP) stages (S1 and S4) (Figure 1A). The mutant showed a similar seed weight to wild type at S1, but displayed a significant decrease in seed weight at S4 (60% of wild-type; Figure 1C). Meanwhile, the mutant exhibited significantly reduced seed length compared with wild-type, and decreased mean seed area at S4 (Figures 1D and 1E). Furthermore, longitudinal sections of representative seeds at S4 were observed (Figures 1F and 1G), and the mutant displayed a much bigger cell area and significantly increased cell width (Figures 1H–1J), suggesting a reduced number of cells in mutant seeds. Therefore, the decrease in both seed size and weight of H1314 is mainly because of a reduction in the number of cells.

Cytosine methylomes of developing seeds from wild type and mutant

To investigate changes in the DNA methylome during peanut seed development, we generated single-base resolution maps of DNA methylation for immature seeds at two different stages (15 and 60 DAP) from the wild type and mutant, respectively. Eight seed samples were sequenced for the two lines with two biological replicates (A and B) for each stage. For each sample, over 299 million clean reads of 150 bp in length were generated, with an average bisulfite conversion rate of 99.3% (Table S1). Approximately 69.3% of the reads were mapped to the Tifrunner (*A. hypogaea*) genome assembly (gnm1), covering ~94.5% of the whole genome. All of the sequenced methylomes had ~19.6×depth per chromosome on average (Table S1). There are 927.4 Mb cytosines in the Tifrunner genome assembly.[19] For the eight analyzed seed methylomes, 7,526 to 8,798 Mb cytosines were sequenced and mapped to the peanut reference genome with an average 8.4-fold depth. Among them, 324−350 Mb cytosines were methylated with the average methylation density of 35.0–37.8% in single seed samples, across CG, CHG and CHH sequence contexts (Figures 2A and 2B). Notably, more than half of methylated Cs are within the CHH context (57.9%), followed by CHG (23.3%) and CG (18.8%; Figure 2C). We then examined methylation levels for sequenced

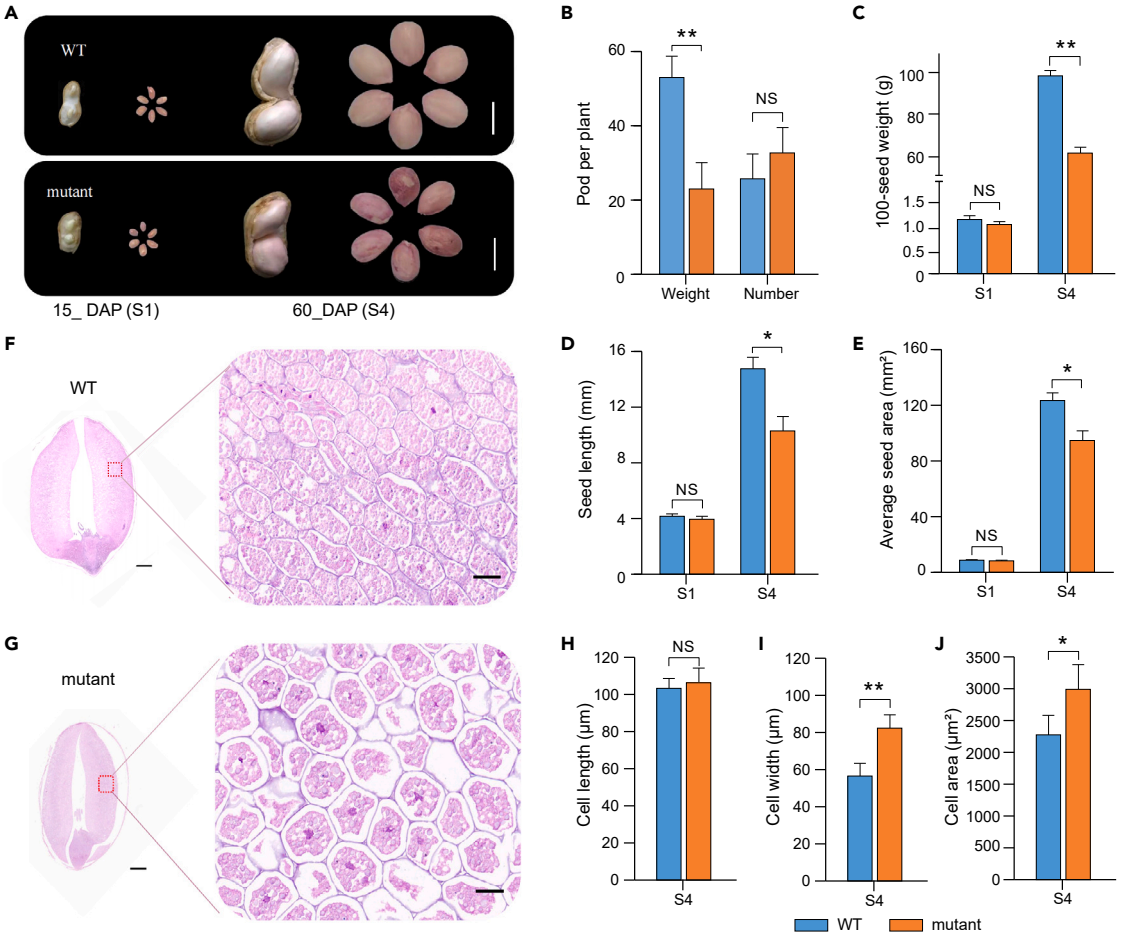


Figure 1. Phenotypes of developing seeds for wild type (H2014) and mutant (H1314) peanut lines

(A) Seed samples at 15 (S1) and 60 DAP (S4) from wild-type (WT, H2014) and mutant (H1314) plants (Bars = 1cm).

(B) Changes in average pod weight and number per plant at maturity, n = 10.

(C–E) Average 100-seed weight, seed length and seed area at both developmental stages. n = 20 in panel C, and n = 10 in panel D and E.

(F, G) Longitudinal sections of wild type seeds at S4 are shown in (F) and those of mutant seeds are shown in (G). Bars = 20 μm.

(H–J) Cell length, width, and area of seeds are displayed in (H), (I) and (J), respectively, for both lines. n = 15. *$p < 0.05$, **$p < 0.01$, NS, not significantly. Student's *t* test. Results are means ± standard deviation (SD).

seed samples, and all pairs of replicates displayed highly positive correlations ($R^2 \geq 0.97$; Figure S1). The global methylation level was ~29.6% across these samples, with average CG, CHG and CHH methylation levels of 79.6%, 69.2% and 15.9%, respectively (Table S2).

We also analyzed the distribution of DNA methylation and gene/transposable element (TE) density. The results revealed enhanced DNA methylation in the three contexts in the pericentromeric region in which TEs are preferentially distributed (Figures 2D and 2F). By contrast, relatively little DNA methylation was observed in the chromosome arms where genes are massively enriched (Figures 2D and 2F). There were lower methylation levels in genic regions, especially 3′/5′ flanking regions, than in 2 kb upstream/downstream regions for both lines (Figures 2E and 2G). Moreover, compared with other regions of the gene body, promoter and intron regions showed enhanced methylation in all three sequence contexts (Figures 2E and 2G).

Elevated methylation during developing peanut seeds of mutant and wild type

Both CG and CHG methylation in seed samples showed strong positive correlations ($R^2 \geq 0.97$) between the two seed stages within wild type or mutant line, whereas reduced coefficients ($R^2 = 0.87–0.90$) were

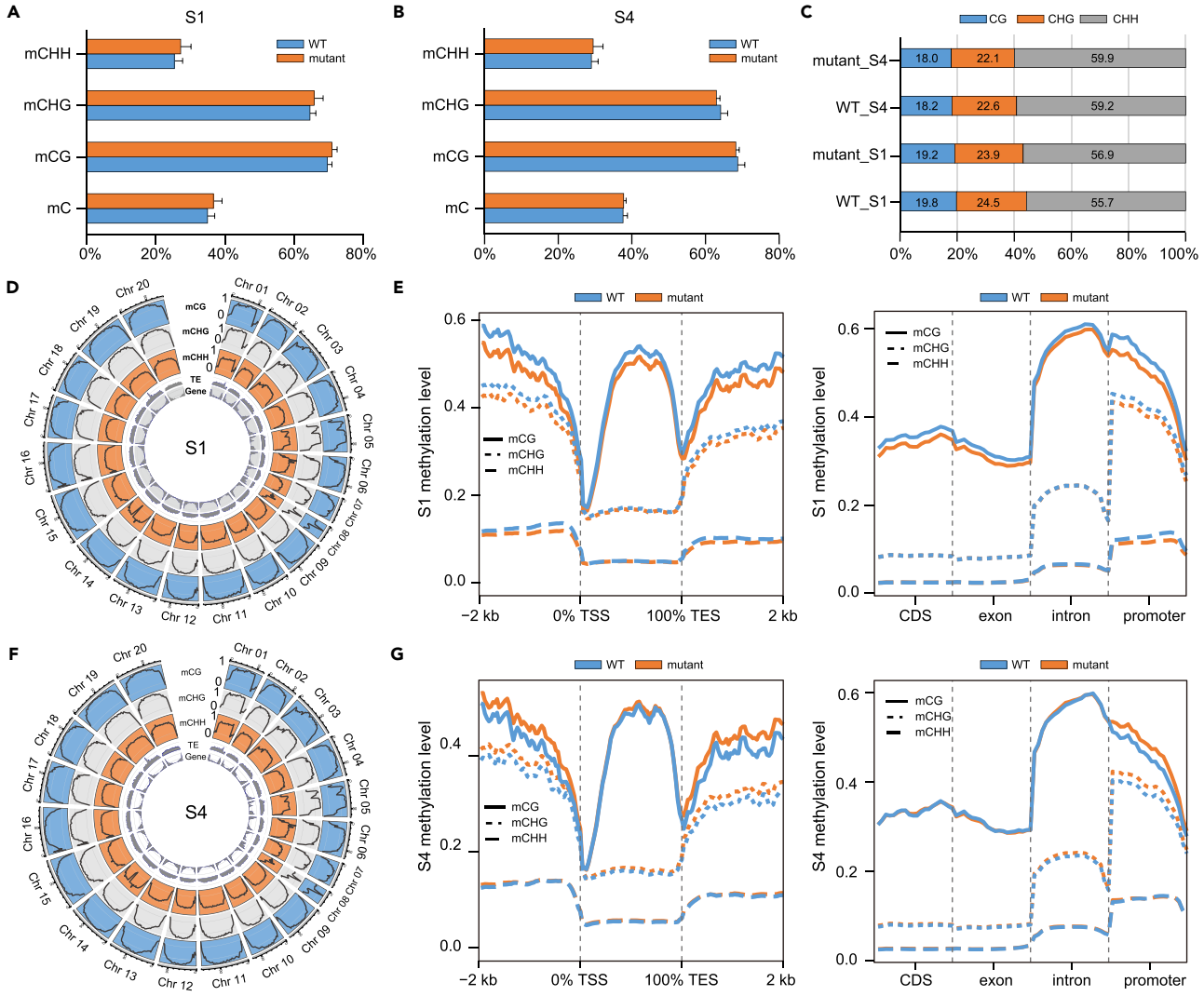


Figure 2. Cytosine methylation profiles of developing seeds of wild type and mutant peanut lines

(A and B) Average density of methylated cytosines (Cs) in wild type (WT) and mutant seeds at 15 (S1, A) and 60 DAP (S4, B). And the percentage for cytosine methylation in CG, CHG, or CHH contexts is shown for both peanut lines. The average methylation density is that the number of identified methylation Cs is divided by total number of sequenced cytosines with coverage depth ≥5, and the percentage was generated for cytosine methylation in CG, CHG, or CHH contexts, respectively.

(C) Percentage of CG, CHG and CHH methylation in single seed sample cytosine methylation profile. The number of cytosines in CG, CHG and CHH contexts was divided by total number of Cs in single seed methylome, and the percentage was generated.

(D and F) Circo heatmaps including cytosine methylation in CG, CHG and CHH contexts (methylation level range, 0–1), TE (transposon elements) and gene intensity (from outer to inner) across the whole peanut genome for wild-type seeds at S1 (D) and S4 (F). TE intensity range, 0–22; gene intensity range, 0–500.

(E and G) Distribution of methylation levels in CG, CHG, and CHH genomic sequence within 2kb up/downstream of gene body, gene body, and functional regions of gene were displayed for wild type and mutant seeds at S1 (E) and S4 (G), respectively. See also Figure S1, Tables S1 and S2.

observed for CHH methylation, suggesting substantial changes in the seed DNA methylomes of mutant and wild type from S1 to S4 (Figure S1). So we examined the dynamic methylomes during peanut seed development. The average methylation level of 28.5% at S1 was globally elevated to 29.9% at S4 in wild type, and a similar trend (from 28.7% to 31.1%) was observed in mutant seeds (Figure 3C; Table S2) in which both peanut lines all had increased methylation levels in promoter, gene body, and repeat at S4 compared to S1 (Figure 3D; Tables S5 and S6). Totally, there were 29,904 and 14,588 differential methylation regions (DMRs) between S1 and S4 in wild type and mutant seed samples, and of them, 27,769 (92.9%) and 13,247 (90.8%) were hyper-DMRs, respectively (Figure 3A; Tables S3 and S4). Of interest, we found that most DMRs and most hyper-DMRs (97.7%, 96.9%) had CHH methylation (Figure 3B). In addition, wild type seeds had

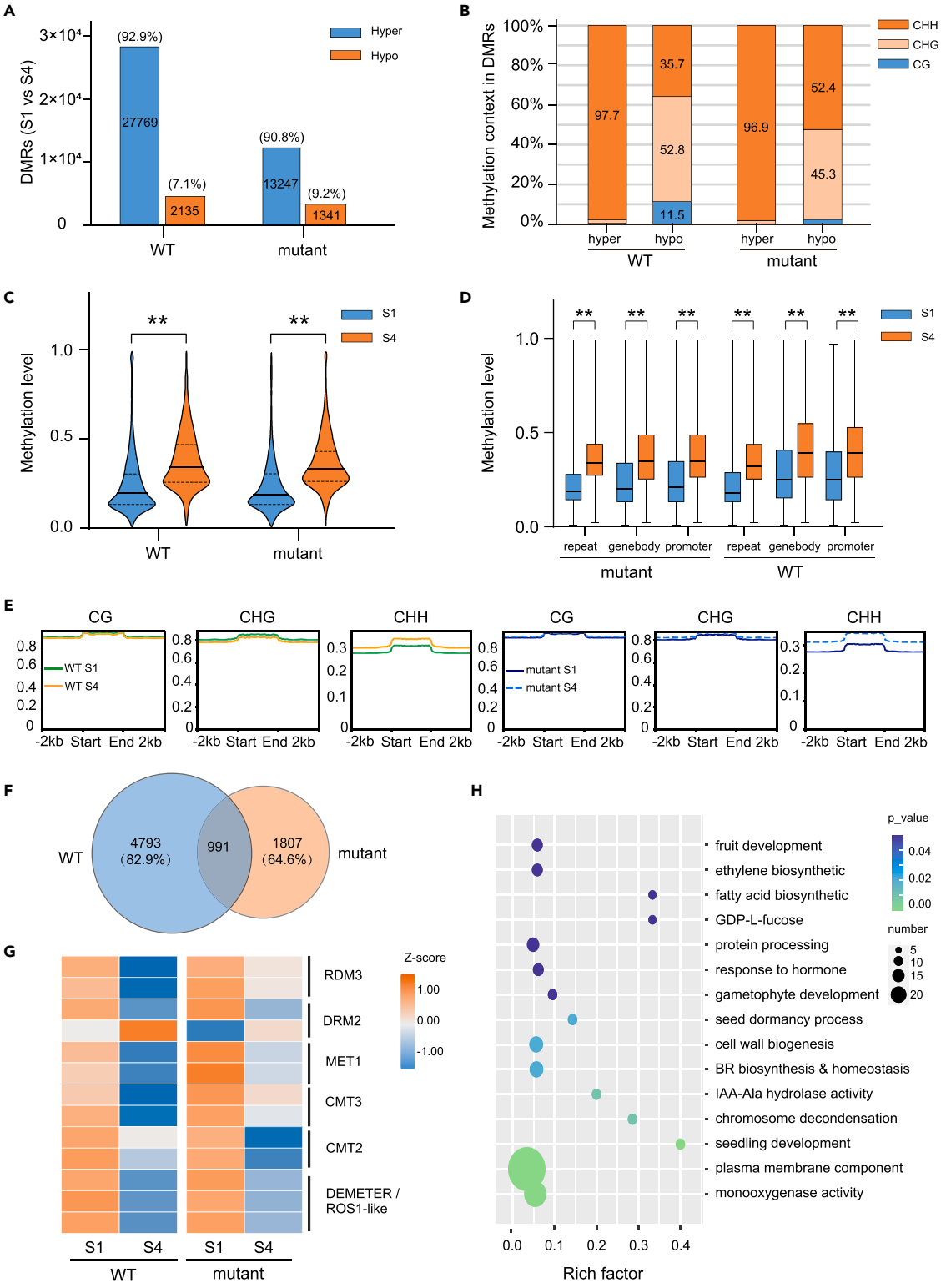


Figure 3. Increased methylation during seed development from S1 to S4 for mutant and wild type peanut lines

(A) The wild type (H2014, WT) or mutant (H1314) seeds at S4 (60 DAP) had differential methylation regions (DMRs) with increased (hyper) or decreased (hypo) cytosine modification compared to S1 (15 DAP).

Figure 3. *Continued*

(B) Hyper-/hypo-DMRs between S1 and S4 for WT or mutant seeds had cytosine methylation in CG, CHG, and CHH genomic context, and the percentages of CG, CHG, and CHH methylation within wild-type or mutant seed DMRs were displayed.

(C) Significantly increased methylation in mutant or wild type seeds from S1 to S4 based on the identified DMRs from panel (A), the total DMRs (including hyper-DMRs and hypo-DMRs) identified in WT or mutant seeds at S4 compared to S1 were used to analyze methylation change during seed development.

(D) DMRs were mostly distributed in repeat, gene bodies, and promoter regions in peanut genome. Differential methylation analyses were performed for the total DMRs (including hyper-DMRs and hypo-DMRs) within repeat, gene body, or promoter identified in wild type or mutant seeds at S4 compared to S1, respectively.

(E) The DNA methylation distribution across the transposable elements (TE) and flanking 2 kb regions in developmental wild type and mutant seeds at S1 and S4.

(F) Venn diagram showing DMRs-targeting genes in developing wild type (blue circle) or mutant (red circle) seeds at S4 compared to S1. The DMRs are the total DMRs (including hyper and hypo-DMRs) within promoter and gene body region, identified in developing wild type or mutant seeds at S4 compared to S1.

(G) Expression heatmap of DNA methyltransferases (*RDM3*, *DRM2*, *MET1*, *CMT3*, *CMT2*) and glycosylases (*DEMETER-like*/*ROS1*) in developmental mutant and wild type seeds (RNA-seq data). For each gene, expression levels (FPKM) from seed RNA-seq data were treated using *Z* score and used to generate the heatmap.

(H) GO enrichment was performed in the common DMR-targeting genes shared by developing wild type and mutant seeds from S1 to S4, and some significantly terms including seed development, gene expression, plant hormone, and so on were present. Student's *t* test, $^{**}p < 0.01$ for (C) and (D). See also Figures S2 and S3, Tables S3, S4, S5, S6, S7, S8, and S16.

5,784 DMRs-targeting genes between S1 and S4, whereas the mutant had 2,798 DMRs-targeting genes, accordingly (Figure 3F). Notably, 991 DMRs-targeting genes were shared by both peanut lines (Figure 3F). We further performed GO enrichment analysis on these DMRs-related genes, and some important terms such as hormone level regulation ("IAA-Ala hydrolase activity", "BR biosynthesis and homeostasis"), lipid & sugar metabolic process, cell wall biogenesis and cell division ("chromosome decondensation") were significantly enriched. More importantly, the terms seed development ("seed dormancy process", "gametophyte development", "seedling development", and "fruit development") were also included, indicating wide involvement and indispensable functions of DNA methylation in peanut seed development processes (Figure 3H; Table S7).

Meanwhile, we have examined the methylation distribution at TE regions from S1 to S4 in developmental wild type and mutant seeds. The result showed that both WT and mutant seeds have similar CG and CHG methylation but significantly increased CHH methylation from S1 to S4, and much higher methylation levels are distributed across TE body than flanking regions (down or upstream 2 kb sequence) (Figure 3E). Transposable elements have been identified in cultivated peanut genome including LTR retrotransposon (Retro), Non-LTR retrotransposon (LINEs and SINEs), and DNA transposons (CACTA, hAT, Helitron, and Mutator).[19] We observed the globally elevated CHH methylation across these transposable elements (TE body and 2 kb flanking regions) from S1 to S4 during developmental WT and mutant seeds (Figure S2).

Elevated cytosine methylation was observed from S1 to S4 for both mutant and wild type seeds, Compared with S1, wild type seeds had obviously enhanced CHH methylation levels but slightly reduced CG and CHG methylation levels at S4, whereas the mutant mainly showed large increases in CHH methylation (Table S2). Therefore, the global increase in methylation levels from S1 to S4 may be mainly attributable to increased cytosine methylation in the CHH context. Methyltransferases and glycosylases are key factors affecting DNA methylation as well as the RNA-directed DNA methylation (RdDM) pathway, and their expression levels directly regulate overall methylation in plant genomes.[22,23] A BLAST search of Arabidopsis methyltransferases and glycosylases against the published Tifrunner (*A. hypogaea*) genome assembly (gnm1) identified 8 glycosylase and 14 methyltransferase homologs in cultivated peanut (Table S8). We further examined expression levels of these genes, and found that ten DNA methyltransferases were downregulated, whereas only the *Domains Rearranged Methyltransferase 2* (*DRM2*)*-like* gene (*AhZ5TSV0*), responsible for *de novo* methylation of cytosines in CG, CHG and CHH sequence contexts through the RdDM pathway, showed increased expression at S4, and the other three were not expressed in developing peanut seeds (Figures 3G and S3; Table S8). We have also examined other key factors in the RdDM pathway, which include RNA-directed RNA polymerase 2 (RDR2) and endoribonuclease DICER LIKE-3 (DCL3) responsible for generation of 24 nt sRNAs, and the Argonaute (AGO) proteins (especially AGO4) necessary for DNA methylation of target loci together with DRM2.[23] And they mostly displayed a marked reduction in expression from S1 to S4 in both peanut lines (Table S8). Meanwhile, of the eight glycosylases including REPRESSOR OF SILENCING 1(ROS1)-like isoform and transcriptional activator DEMETER-like protein, five displayed decreased expression levels, and two were unexpressed (Figures 3G and S3; Table S8).

Thus, it seems that both DRM2-like (AhZ5TSV0) and the five glycosylases with reduced expression could collectively contribute to globally elevated methylation during seed development in the two peanut lines.

Differential methylation between developing mutant and wild type seeds

We next compared the mutant and wild type seed methylomes. At S1, the mutant had lower DNA methylation levels in all three sequence contexts (CG, CHG and CHH) at 2 kb upstream/downstream of genebody regions compared with wild type, whereas similar CHG and CHH methylation levels were observed in genic regions for both peanut lines (Figure 2E). At S4, the mutant had enhanced DNA methylation in CG and CHG sequence contexts than wild type, but similar CHH methylation levels (Figure 2G). The methylated cytosines in genic regions were further examined, and the results showed that at S1, the mutant seeds had globally reduced DNA methylation in the CG context across gene body regions, and decreased methylation levels in CHG and CHH contexts in promoter regions (Figure 2E). At S4, both CG and CHG methylation levels in promoters were obviously elevated in the mutant, whereas DNA methylation in other genic regions was similar to the wild type parent (Figure 2G).

We then characterized DMRs among seed methylomes, and identified 19,833 and 29,023 DMRs between mutant and wild type seeds at S1 and S4, with an average fragment length of 268 bp and 247 bp, respectively (Figured S4B and S4C; Tables S9 and S10). Among them, the mutant had many more hyper-DMRs (10,162) at S1 but more hypo-DMRs (16,699) at S4, corresponding to 51.2% and 57.5% of total DMRs, respectively (Figure S4A). At S1 stage, more DMRs with CG (34.1%, 6,759) and CHG (34.4%, 6,833) methylation were present than CHH (31.5%, 6,241) methylation (Figure S4B), whereas CHH (49.2%, 14,291) methylation were predominant followed by CG (27.3%, 7,913) and CHG (23.5%, 6,819) methylation at S4 stage (Figure S4C). We have also identified the DMRs specifically at S1 or S4, and the common DMRs at both stages. Totally, 42.4–51.3% of the CG and CHG DMRs identified at S4 are shared by both stages, whereas only 4.9–6.8% of the CHH DMRs in mutant seeds at S4 are present at S1 (Figure S5A), indicating that most CHH DMRs at S4 are attribute to accumulated changes from S1 to S4. In addition, the identified hypo-/hyper-DMRs were mainly located in intergenic regions (64.1–72.9%) whereas only a small number (27.1–35.9%) were distributed in promoter and/or genic regions of mutant at both stages (Figure 4B). Importantly, the mutant had many more hyper-DMRs with CG and CHG methylation but fewer hyper-DMRs in CHH sequence contexts at both stages (Figure 4A), which was consistent with increased CG and CHG methylation but reduced CHH methylation at both stages compared to wild type (Figures 4C, 4D, and S5B). Compared with DMRs in intergenic regions, methylation variation in genic regions and promoters may be more important for gene expression. Thus, we further analyzed the hypo-/hyper-DMR distributions in promoters and other genic regions. Totally, mutant seeds had many more hypo-DMRs than hyper-DMRs in all promoter, intron, exon, transcription start site (TSS) and transcription end site (TES) regions at both stages (Table S11). Notably, promoter regions contained 2,526 and 4,160 hyper-/hypo-DMRs at S1 and S4 stages in the mutant, whereas 6,800 and 7,904 DMRs were distributed within the gene body region, respectively (Table S11). Further analysis revealed that 7,548 and 3,688 genes were targeted by the hypo-/hyper-DMRs located in promoters at S1 and S4, respectively, whereas the hypo-/hyper-DMRs distributed in gene body regions targeted 7,008 and 4,393 genes (Figures 4E and S4D; Tables S12 and S13).

The mutant line (H1314) was generated after chemical treatment of the peanut accession H2014, and substantially differential cytosine methylation existed among mutant and wild-type seed samples. To examine whether sequence variation resulting from EMS treatment may affect methylation, we performed sequence analysis using RNA sequencing (RNA-seq) data. A total of 20,276 single base variations were identified in the mutant genome, of which C>T (9,114) and G>A (8,597) were predominant, whereas other variation types were rarely detected (Table S14). Notably, 9,794 Cs were converted to A, T or G, but only 713 cytosines resulted from conversion of non-C bases, indicating a substantial decrease in cytosines in the H1314 EMS mutant. Therefore, elevated methylation in mutant seeds was not directly related to reduction of cytosines in the genome.

Differential gene expression between wild type and mutant seeds

To further investigate the correlation between DNA methylation and gene expression during peanut seed development, we sequenced eight RNA samples extracted from developing seeds of mutant and wild type (S1 and S4) with two biological replicates (A and B) for each stage (Table S15). Pearson correlation analyses revealed a strong positive correlation ($R^2 \geq 0.93$) among seed transcriptomes at the same stage between mutant and wild type, indicating similar gene expression profiles in mutant and wild type seeds

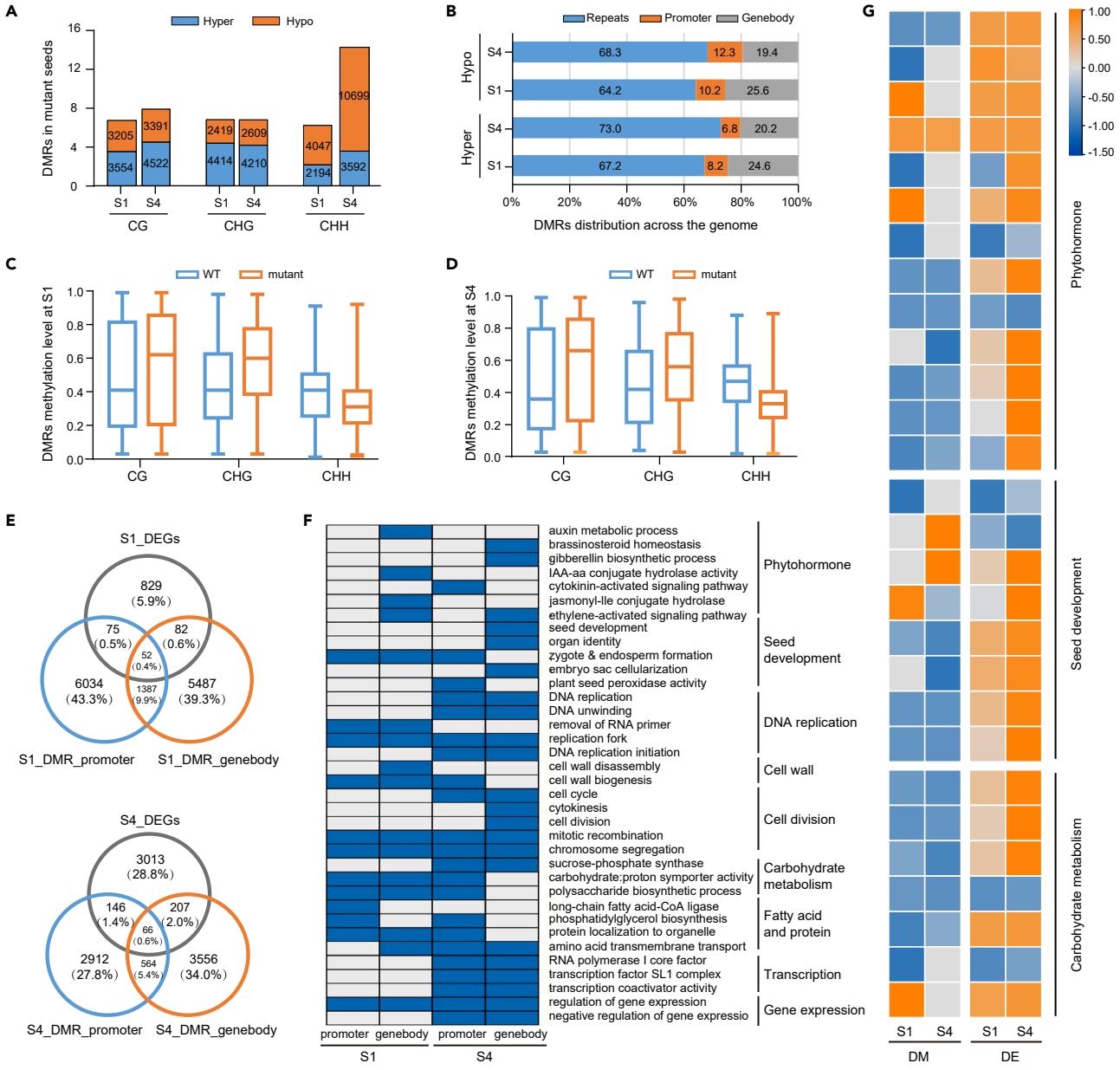


Figure 4. Differential cytosine methylation between the wild type and mutant developing seeds

(A) Differentially methylated regions (DMRs) with CG, CHG, and CHH methylation between mutant and wild type seeds were shown at both stages. Hyper-/hypo-DMRs, DMRs with increased or decreased methylation, respectively. S1/S4, 15/60 DAP.

(B) Distribution of hyper-/hypo-DMRs in promoter, gene body, or other region (repeat), identified in mutant at both stages when compared to wild types.

(C and D) Comparison of methylation level in CG, CHG, and CHH contexts within the identified DMRs from panel (A) between mutant and wild type seeds at S1 (C) and S4 (D).

(E) Venn diagram showing differentially expressed genes (DEGs) and DMRs-targeting genes with differential methylation in promoter (DMR_promoter) and/or gene body (DMR_genebody) identified in developmental mutant seeds when compared to wild type samples at both stages.

(F) GO enrichment was performed on the DEGs with differential methylation in promoter and/or gene body between mutant seeds and wild type samples at S1 and S4 stages. S1-/S4-promoter, the DEGs showing differential methylation within promoter at S1/S4; S1-/S4-genebody, the DEGs showing differential methylation within gene body at S1/S4; the blue box, significantly enriched terms with $p \leq 0.05$ whereas for the white box $p > 0.05$.

(G) Some genes related to the terms "seed development, hormone, and carbohydrate metabolism" from panel (F), and the corresponding differential methylation (DM) and expression (DE) between mutant and wild type at both stages. DM = mutant-wild type; DE, log2Fold Change, Fold Change of expression abundance = mutant/wild type; the data were normalized using *Z* score. See also Figures S4–S6, Tables S9, S10, S11, S12, S13, S16, S17, S18, S21, S22, and S23.

(Figure S6A). Approximately 32,000−36,000 gene loci were expressed (FPKM ≥1) in each of the eight seed samples. Compared with S1, the number of low-expression genes (1 ≤FPKM <3) was slightly increased (from 9,159 to 9,870 in wild type) in both lines at S4 stage, whereas high-expression genes (FPKM ≥15) were markedly reduced (from 7,116 to 4,915 in wild type seeds; Table S16).

Compared to wild type, the mutant had 1,038 and 3,432 DEGs at S1 and S4, respectively, of which most were upregulated at the latter stage of seed development (Tables S17 and S18). Furthermore, 288 commonly upregulated and 141 commonly downregulated genes were present in the mutant at both stages (Figure S6B). The commonly upregulated genes were enriched in the pathways such as "mismatch repair", "amino sugar and nucleotide sugar metabolism", "MAPK signaling pathway-plant", "pentose and glucuronate interconversions", and "glycolysis/gluconeogenesis"; meanwhile, the shared downregulated genes were mostly enriched in "starch and sucrose metabolism", "galactose metabolism", "phenylalanine, tyrosine and tryptophan biosynthesis", and "diterpenoid biosynthesis" (Figure S6C; Tables S19 and S20). Notably, the carbohydrate metabolism-associated genes showed down-regulation in mutant seed samples at S1 and S4, whereas the genes related to hormone signal transduction were up-regulated at both stages of seed development, which might be responsible for the reduced seed size and weight of the mutant line (Figures S6C and S6D).

Joint analyses of seed transcriptomes and methylomes of mutant and wild type

Based on the seed transcriptomes and methylomes generated above, we next examined the correlation between gene expression and cytosine methylation. In general, genes located on chromosomal arms had high expression levels but showed reduced methylation, however, those distributed in pericentromeric regions always displayed low FPKM values but had enhanced cytosine methylation, suggesting a negative correlation between gene expression and cytosine methylation overall (Figures 2D, 2F, S7A, and S7B). In addition, we found that cytosine methylation in different sequence contexts had various effects on gene expression levels. Compared with genes not expressed (FPKM <1), expressed genes (1≤ FPKM) showed globally increased methylation in the CG context within 2 kb upstream, but they displayed an obvious reduction in mCG at the 5′/3′ ends of gene body regions (Figure S7C). In addition, the expressed genes generally had lower methylation levels in both CHG and CHH contexts across the entire gene body regions. Moreover, with increasing expression levels, methylation levels in CHH within 2 kb upstream were elevated during development in seeds of both lines (Figure S7C). In addition, we also examined the correlation between gene expression and cytosine methylation among different genetic background. Some genes (e.g., *AhFMLD33* and *AhZ6Z4XE*) showed increased methylation and elevated expression levels in mutant and wild type seeds, whereas others (e.g., *AhK8H9R8* and *AhAXL13P*) had enhanced methylation but decreased expression, suggesting complicated relation between transcription levels and DNA methylation (Figure S4F).

We further examined whether DNA methylation may alter expression levels of the corresponding genes. There were totally 1,038 and 3,432 DEGs between mutant and wild type at S1 and S4, respectively (Figure S6B; Tables S17 and S18). Of them, the mutant seeds had 209 DEGs with DMRs (DMR-DEGs) in promoter (127) and gene body (134) at S1, and 419 DMR-DEGs in promoter (212) and gene body (273) at S4 (Figures 4E and S4D). GO enrichment were further performed on these DMR-DEGs. In summary, the terms like "phyhormone", "seed development", "DNA replication", "cell division", "transcription", and "gene expression" were over-represented (Figure 4F; Table S21). Notably, GO terms "auxin metabolic process & homeostasis" were significantly enriched at S1, whereas "gibberellin biosynthetic process", "brassinosteroid homeostasis", and "cytokinin-activated signaling pathway" were mainly enriched at S4. In addition, some DMR-DEGs were significantly enriched in the terms "seed development", "organ identity" and "embryo sac cellularization" at S4 stage, which was possibly associated with large seed size variations between mutant and wild type (Figures 4F and 4G; Table S21).

The mutant showed a similar seed weight to wild type at S1, but displayed a significant decrease by 40% at S4. To explore further the key genes determining seed phenotype variation differentiating mutant from wild type, we mainly analyzed the DMR-DEGs identified in mutant seeds at S4. Totally, the mutant seeds had 212 and 273 DMR-DEGs within promoter and gene body at S4, and of them, 118 and 168 showed negative correlation between expression fold change and methylation variation, respectively, suggesting the important effects of DNA methylation on gene expression regulation and potential roles of these DMR-DEGs during peanut seed development (Figures 5A and 5B; Tables S22 and S23). Furthermore, 38 genes

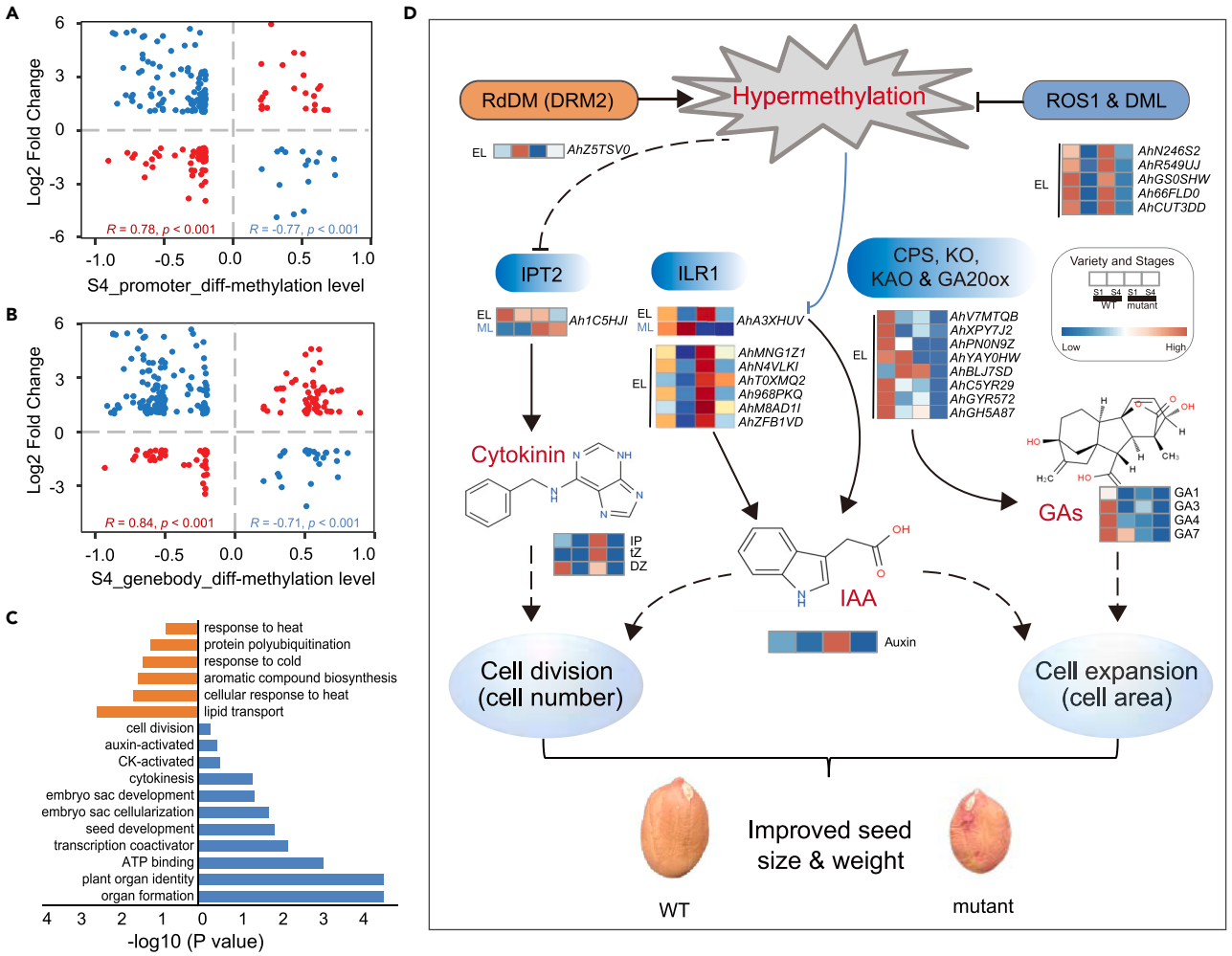


Figure 5. Differentially expressed and methylated genes between the wild type and mutant developing seeds (60 DAP) associated with peanut seed development

(A and B) Correlation between gene expression and cytosine methylation within promoter (A) or gene body (B) at 60 DAP (S4). The developmental mutant seeds had differential methylation-targeting genes with differential expression compared to wild type samples at S4, and on these genes Pearson's correlation analysis was performed.

(C) The genes showing negative correlation between differential mRNA abundance and differential methylation distributed in promoter (118 genes, A) or gene body (168 genes, B) were selected, and GO enrichment was carried out. The terms included phyhormone (auxin, cytokinin, and GA), cell division, and reproduction organism development. Orange bars, enriched terms on the genes that were downregulated but had significantly increased methylation within promoter and/or gene body at S4; blue bars, enriched terms on the genes that were upregulated but had significantly decreased methylation at S4; CK/auxin-activated, cytokinin-activated signaling pathway.

(D) Putative model for DNA cytosine methylation affecting peanut seed size/weight by altering endogenous cytokinin and auxin levels. The *DRM2-like* gene (*AhZ5TSV0*), a peanut DNA methyltransferase involved in *de novo* methylation of cytosines; homologs of Arabidopsis *ROS1-like* (*AhN246S2*, *AhR549UJ*, *AhGS0SHW*) and *DEMETER-like* (*Ah66FLD0*, *AhCUT3DD*). EL, expression level (black font); ML, methylation level (blue font); solid lines indicate pathways confirmed or identified in previous studies; dashed lines indicate unconfirmed or speculative pathways. Heatmaps next to "Cytokinin, IAA, and GAs" represented the corresponding hormone levels. *IPT2* (*Ah1C5HJI*), peanut tRNA *isopentenyltransferase 2*, an Arabidopsis *IPT2* homolog, is involved in positive regulation of CK biosynthesis. IAA-amino acid hydrolase *ILR1-like* genes (*ILR1*) including *AhA3XHUV*, *AhA3CFJS*, *AhMNG1Z1*, *AhT0XMQ2* and *AhN4VLKI*. GA biosynthesis-related genes: ent-kaurenoic acid oxidase 2 (*KAO2*), *AhV7MTQB*; gibberellin 20 oxidase 2 (*GA20ox2*), *AhBLJ7SD*, *AhC5YR29* and *AhGYR572*; ent-kaurene oxidase (*KO*), *AhXPY7J2*; ent-copalyl diphosphate synthase (*CPS*), *AhPN0N9Z* and *AhYAY0HW*. Gene expression levels (FPKM, EL) from seed RNA-seq data and methylation levels (ML) were normalized using *Z* score to generate the expression heatmap or methylation heatmap, respectively. Heatmaps next to the chemical formula for "Cytokinin, IAA, and GAs" are the corresponding endogenous hormone levels from fresh wild type or mutant seeds at S1 and S4, in which the values were treated using *Z* score. See also Figures S6–S8, Tables S16, S17, S18, S24, S26, S27, S28, and S29.

were downregulated but had significantly increased methylation at S4, and on these genes the terms "lipid transport", "transmembrane transporter", "aromatic compound biosynthesis", and "methyltransferase activity", were significantly enriched (Figure 5C; Table S24). Meanwhile, the mutant seeds had 202 upregulated genes but significantly decreased methylation at S4, and on these DMR-DEGs some important terms related to seed development were over-represented like "plant organ identity", "sucrose biosynthetic process", "seed development", "embryo sac development" (Figure 5C; Table S24). Notably, the GO terms "gibberellin biosynthetic process", "brassinosteroid homeostasis", "cytokinin and auxin-activated signaling pathway" were included (Figure 5C; Table S24). Of interest, some hormone-related pathways terms like "zeatin biosynthesis", "plant hormone signal transduction", "diterpenoid biosynthesis", and "auxin transport" were also concurrently enriched in the identified DEGs (Figures S6C and S6D; Table S25) between mutant and wild type samples, suggesting the potential roles of the identified DMR-DEGs in peanut seed development.

Association of methylation variation with expression of key genes involved in hormone metabolism

We further examined cytosine methylation and expression levels of these hormone metabolism-related genes. CKs play important roles in promoting cell division.[5] A total of 112 CK biosynthesis and signaling-related genes were obtained from the genome assembly of the segmental allotetraploid peanut *A. hypogaea*,[19] and 21 with expression levels of FPKM $\geq$2 were used for further methylation examination (Table S26). In developing mutant seeds at S1 and S4, 10 genes had elevated or reduced CG/CHH methylation in promoter or intron/exon regions, whereas only four genes had differential cytosine methylation at both stages. Previous studies have shown that the ATP/ADP-isopentenyltransferase (*AtIPT*) gene family are among key enzyme families catalyzing the rate-limiting step of CK biosynthesis in Arabidopsis[5]. Notably, *Ah1C5HJI*, encoding tRNA isopentenyltransferase 2, the *Arabidopsis IPT2* homolog, showed consistently elevated CG methylation in the promoter at both stages in mutant, but displayed reduced expression levels (Figure 5D; Table S26). Meanwhile, we performed sequence analysis, and no DNA sequence variation was observed between mutant and wild type, suggesting that increased CG methylation may affect transcription levels of *Ah1C5HJI*.

In view of the concurrent enrichment of "auxin transport or auxin-activated signaling pathway" and "diterpenoid biosynthesis" in both DMR-targeting DEGs and the identified DEGs in the mutant line, methylation analysis was also conducted on auxin and gibberellin (GA) metabolism-related genes. Of 59 auxin biosynthesis and signaling-associated genes with expression of FPKM $\geq$2, 41 had increased or decreased CG/CHH methylation in promoter or intron/exon regions at S1 and/or S4, and 21 had differential methylation at both stages, of which reduced methylation was predominant (Table S27). Specifically, the IAA-amino acid hydrolase *ILR1-like 4* (*AhA3XHUV*), a close homolog of *Arabidopsis ILL4*, involved in regulating free IAA accumulation in plants,[24] showed reduced CHH methylation in the promoter at the early stage, whereas transcript levels were upregulated (Figures 5D and S8A–S8D; Table S27). Meanwhile, the mutant seeds had increased CG methylation within the intron of the peanut IAA-amino acid hydrolase *ILR1-like 6* (*AhT0XMQ2*, a close homolog of *Arabidopsis ILL6*) at S1, but decreased CHH methylation in the intron at S4; however, the gene was upregulated in H1314 seeds at both stages (Figures S8A–S8C; Table S27). Of interest, *AhTR1S4A*, encoding for serine/threonine-protein kinase *BRI1-like 2* (*BRL2*), another gene involved in auxin-activated signaling pathway (GO:0009734, Table S24), had markedly reduced CG and CHG methylation in the promoter in mutant seeds at S1, and slight decrease in CHH methylation was also present across the gene body and promoter at S4; whereas the expression was upregulated in developmental mutant seeds at both stages (Figure S8A; Table S27). In addition, other auxin signaling related-genes such as auxin-responsive protein *IAA9* (*AhXR81AR*) had reduced CG methylation of intron/exon regions, whereas pyrophosphate-energized membrane proton pump 2 (*AhAK3WVB*) showed a stable increase in CG methylation in the promoter (Table S27). Regardless of whether methylation was elevated or reduced among these genes, there seemed to be an opposite trend between expression levels and methylation levels (Figure S8A). It also indicated important roles of DNA cytosine methylation in auxin level regulation and signaling pathway. Meanwhile, in the mutant seeds 22 GA biosynthesis and transduction-related genes, including ent-kaurenoic acid oxidase (*KAO*), GA20 oxidase (*GA20ox*), ent-copalyl diphosphate synthase (*CPS*) and GA3-beta-dioxygenase (*GA3ox*; Table S28), had differential cytosine methylation, and they mostly displayed decreased CG/CHG/CHH methylation in promoter, exon, and/or intron regions.

Reduced CK and GA levels and elevated IAA content contribute to smaller mutant seeds

To investigate the cause of variation in seed size/weight between mutant and wild type, we also measured CK, IAA and GA levels in developing seeds of both mutant and the wild type at S1 and S4. For CKs, three (iP, tZ and DZ) of four types were detected (cZ was not), DZ was predominant, and CK levels displayed a globally significant decrease in both peanut lines from S1 to S4 (Figure 5D; Table S29). Notably, mutant seeds had much higher iP and tZ levels at S1, but showed a significant reduction in DZ content at both stages, 72% and 80% of levels in wild type seeds, respectively, indicating a marked decrease in total CK levels in developing mutant seeds (Figure 5D; Table S29). Unexpectedly, IAA levels in mutant seeds were increased significantly (1.5-fold) at S1, but reduced slightly by 24% at S4 (Figures 5D and S8D; Table S29). Meanwhile, mutant seeds showed a significant decrease in both bioactive GAs and inactive precursor/synthetic intermediates, except for GA_5 and GA_{20}, at both stages, consistent with the reduced gene expression levels explained above. Notably, compared with those in wild-type samples, levels of GAs with bioactivity (GA_3, GA_4 and GA_7) in mutant seeds were decreased >3-fold at the two stages, whereas GA_1 was not detected at S4 stage in both peanut lines (Figure 5D; Table S29).

As mentioned above, mutant seeds showed significantly reduced seed length but much larger cell area, indicating a reduced number of cells (Figures 1D and 1J). It is well known that CKs play important roles in promoting cell division.[5] Therefore, the reduced mutant seed size/weight may be mainly attributed to significantly decreased total CK levels, especially DZ, as well as the coordination of enhanced IAA and decreased GA content.

DISCUSSION

Peanut bears pods or seeds depending on gynophore or peg formation, a special tissue arising from a pollinated flower that elongates to send the ovary or embryo into soil for pod enlargement and seed development.[25] In previous studies, either small amounts of tip tissue of developing pegs, a mixture of maternal and zygotic tissues,[26] or young unexpanded leaves,[19] were collected for DNA methylation examination, hence the results could not reflect actual methylation profiling during peanut seed development. By comparison, here we characterized seed methylomes at early and late stages of development, hence our data could be used to explore epigenetic regulation related to peanut seed development. In addition, DNA methylation varies across different plant species. In our study, developmental peanut seeds had a global methylation level of ~29.6% (Table S2), much higher than observed for seeds of soybean (~12%) and *Arabidopsis thaliana* (~5.5%),[27] fruits of tomato (~22%)[28] and orange (~13%).[17] It is noteworthy that global methylation level is elevated in developmental peanut seeds at S4 (late maturation stage) compared to S1, whereas the relatively constant methylation is present among whole soybean seeds at various stages including the globular, cotyledon, early maturation, mid maturation, late maturation stages, or among developmental Arabidopsis seeds at globular, linear cotyledon, mature green, postmature green stages.[27] Of interest, increased CHH methylation levels are observed in developmental soybean, Arabidopsis, and peanut seeds, suggesting a probably conservative trend of CHH change in dicotyledons, which needs to be confirmed. DNA methylation is generally enriched in heterochromatin or pericentromeric regions where TEs are preferentially located,[17,27] hence the observed increase in methylation in peanut may be associated with increased TE levels in the peanut genome.[19]

Dynamic cytosine methylation is modulated in a coordinated manner by methylation and demethylation reactions through different mechanisms.[29] Functional mutation of the tomato *SlDML2* gene, a homolog of *Arabidopsis* DNA demethylase *ROS1*, resulted in a genome-wide increase in DNA methylation (hypermethylation) during fruit ripening.[30] A recent study combining reverse and forward genetics approaches demonstrated that DNA methyltransferases are responsible for overall DNA methylation in *Arabidopsis*.[22] It is well known that RdDM is responsible for *de novo* establishment of DNA methylation in plants, and functional defects in rice RNA-dependent RNA polymerase 2 (*OsRDR2*) caused a global decrease in CHH methylation by inhibiting the accumulation of 24 nt small interfering RNAs.[31] Notably, in the present study, immature peanut seeds displayed globally elevated methylation from early to late developmental stages and a substantial increase in CHH methylation, which might be also mainly attributed to reduced expression of most *ROS1* and *DEMETER-like* genes, indicating important parts of demethylases in methylation regulation.

DNA methylation affects gene expression. In plant genomes, gene distribution is negatively associated with cytosine methylation.[15,17,27,32] Similar results were also observed in the present study. In addition,

we found that the expressed genes generally showed a notable decrease in cytosine methylation in CHG and CHH contexts across the entire gene body, as well as reduced CG methylation at the TSS and TES (Figure S7C). There appears to be a globally negative correlation between expression level and DNA methylation level to some extent, and this has been studied in crops. Hypermethylation in the promoter of *TaGli-c-2.1*, a member of the γ-gliadin multigene family, which are negative regulators of wheat quality, results in reduced expression and greatly improved gluten strength in bread wheat.[33] Wang et al. also reported that methylation of the promoter of *Wdreb2*, a wheat dehydration-responsive element-binding protein, is negatively correlated with expression.[34] In our study, large numbers of DMR-targeting genes were identified in the smaller mutant seeds at both developmental stages, including some key genes controlling CK and auxin levels. Among them, promoter of peanut *tRNA isopentenyltransferase 2* (*Ah1C5HJI*), a homolog to Arabidopsis *IPT2,* a key enzyme regulating CK biosynthesis, showed significantly increased CG methylation at both stages, but exhibited reduced expression, suggesting a repression effect of DNA methylation on gene expression (Figure 5D; Table S26). A similar phenomenon occurred with some auxin metabolism-related genes including Arabidopsis *ILL1* homolog in peanut (*AhA3XHUV*) and auxin-responsive protein IAA9 (*AhXR81AR*), which displayed reduced CHH methylation in the promoter or decreased CG methylation in intron/exon regions in developing mutant seeds, but increased transcription levels (Figure S8A; Table S27). However, other genes such as peanut *auxin response factor 9* (*AhWDRD4S*), an auxin signaling-related gene, behaved differently. Mutant seeds had slightly decreased CHH methylation in the promoter of *auxin response factor 9* (*ARF9*), and reduced gene expression at both development stages, indicating the potentially enhancing effect of methylated cytosines on gene expression (Table S27). Of interest, levels of Arabidopsis *AtROS1* transcripts can be up- and down-regulated by promoter hypermethylation or hypomethylation, respectively.[35] Furthermore, reduced DNA methylation caused by tomato *SlDML2* (5-^{m}C DNA demethylase) is associated with both the repression and activation of numerous fruit ripening-related genes in tomato.[30] Therefore, DNA methylation may exert various regulatory effects on gene expression during peanut seed development.

Combined seed phenotype and cytological analysis revealed that reduced cell number mainly decreased seed length and weight of mutant (Figures 1D and 1J). Plant hormones play vital roles in seed development, and auxin and CKs have been implicated in determining seed size and weight by regulating cell division and cell proliferation.[7,10] As the predominant CK in peanut seeds, the mutant had significantly reduced DZ levels at the early stage of seed development (Figure 5D; Table S29). Adenosine phosphate-isopentenyltransferase (IPT) is one of two key factors catalyzing CK biosynthesis in plants.[36] Of the nine homologs of Arabidopsis *IPT* in the cultivated peanut reference genome,[19] only penaut *AhIPT2* (*Ah1C5HJI*) was highly expressed in developing peanut seeds (Figure 5D). Similarly, *ZmIPT2* is strongly expressed in developing maize kernels, which coincides with elevated CK levels and occurrence of cell division.[37] Herein, *AhIPT2* showed a notable decrease in expression in mutant seeds at both stages. Therefore, the decrease in CKs (DZ) may result from reduced *AhIPT2* transcript levels in peanut.

Auxin influences cell division, cell elongation and cell differentiation.[38] Of interest, IAA, a typical auxin, was significantly elevated (2-fold) at the early stage of seed development in the mutant (Figure 5D; Table S29), consistent with the enlarged cell area. Endogenous free IAA levels are mainly determined by *de novo* biosynthesis from tryptophan (Trp)-independent and Trp-dependent pathways, but are also regulated by de-conjugation of auxin conjugates (inactive auxin storage form) by IAA-amino acid hydrolyses.[39] For example, a large decrease in free IAA content, caused by mutational loss of expression of the *ZmYuc1* gene encoding indole-3-pyruvate (IPA) monooxygenase YUCCA1, resulted in a marked reduction in dry mass of maize endosperm.[40] Also, functional defects in the IAA-glucose hydrolase TGW6 caused reduced IAA levels and affected rice grain weight.[9] In the present study, many auxin biosynthesis-related genes, such as *YUCCA10* encoding an IPA monooxygenase, were strongly expressed or significantly upregulated in developing mutant seeds (Table S27). Arabidopsis YUC family homologs catalyze the conversion of IPA to IAA in the major auxin biosynthesis pathway (the IPA pathway).[41] Additionally, most peanut IAA-amino acid hydrolases regulating de-conjugation of auxin conjugates were upregulated in immature mutant seeds (Figures 5D and S8B; Table S27). Strikingly, among them, the IAA-amino acid hydrolase *ILR1-like 4* (*AhA3XHUV*) showed a notable decrease in CHH methylation in its promoter in the mutant, and the opposite trend was observed between expression level and cytosine methylation level. Thus, upregulation of both auxin biosynthesis and de-conjugation-related genes may jointly contribute to elevated IAA levels in mutant seeds.

Auxins and GAs can modulate cell expansion synergistically during fruit development,[42] and auxins can enhance GA levels.[43] Surprisingly, unlike the observed increase in auxin, the mutant showed a significant reduction in active GAs at both stages of seed development, but it had enlarged cells (Figure 5D; Table S29), suggesting a major role for auxin in regulation of cell expansion. Recently, CK-auxin interactions were found to play important roles in plant development. Strikingly, antagonistic interactions between CKs and auxins have been implicated in modulation of root meristem size.[44] Of interest, in the present study the mutant contained elevated free IAA levels and enlarged cells, but reduced CK levels and a notable decrease in the total number of cells in seeds, suggesting interactions between CKs and auxins in the regulation of seed size and weight, but this needs further research. Overall, a significant decrease in CKs and increased auxin levels jointly led to a substantial reduction in seed size and weight in the mutant. Based on the results, we propose a schematic model (Figure 5D) in which variation in DNA methylation could regulate peanut seed size and weight by altering the expression of auxin and CK metabolism-related genes.

Limitations of the study

We generated single-base resolution maps of DNA methylation for immature seeds at two different stages of development for wild type peanut line and its small seed mutant, identified important methylation variations associated with seed development, and analyzed the correlations between cytosine methylation and transcription levels. However, more work is needed to confirm the effects of methylation on gene expression and seed development in peanut.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Plant materials
- METHOD DETAILS
  - Endogenous phytohormone determination
  - Cytological observation
  - Whole-genome bisulphite sequencing data
  - RNA-seq, data processing and gene annotation
  - DEGs analyses, GO and KEGG enrichment
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2023.107062.

ACKNOWLEDGMENTS

This work was supported by Grants from Key program of National Natural Science Foundation of China (NSFC)-Henan United Fund (No. U22A20475; U1704232) and Key Scientific and Technological Project in Henan Province (Grant No. 221111110500; 201300111000; HARS-22-05-G1).

AUTHOR CONTRIBUTIONS

Y.D. conceived and supervised the project. Z.F.L., Q.L., K.Z., and Y.Z. analyzed the data. D.C., Z.C., K.K.Z., R.R., and D.Q. performed research. X.Z., X.M., and F.G. assisted in editing the manuscript. Q.M., G.Z., S.H., Z.L., and K.W. collected samples, managed field research and plant propagation. Z.F.L. wrote the draft. Y.D. and Y.Z. revised the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: February 23, 2023
Revised: April 21, 2023
Accepted: June 2, 2023
Published: June 7, 2023

REFERENCES

1. Lu, Q., Liu, H., Hong, Y., Li, H., Liu, H., Li, X., Wen, S., Zhou, G., Li, S., Chen, X., and Liang, X. (2018). Consensus map integration and QTL meta-analysis narrowed a locus for yield traits to 0.7 cM and refined a region for late leaf spot resistance traits to 0.38 cM on linkage group A05 in peanut (*Arachis hypogaea* L.). BMC Genom. *19*, 887. https://doi.org/10.1186/s12864-018-5288-3.

2. Zhang, S., Hu, X., Miao, H., Chu, Y., Cui, F., Yang, W., Wang, C., Shen, Y., Xu, T., Zhao, L., et al. (2019). QTL identification for seed weight and size based on a high-density SLAF-seq genetic map in peanut (*Arachis hypogaea* L.). BMC Plant Biol. *19*, 537. https://doi.org/10.1186/s12870-019-2164-5.

3. Gangurde, S.S., Wang, H., Yaduru, S., Pandey, M.K., Fountain, J.C., Chu, Y., Isleib, T., Holbrook, C.C., Xavier, A., Culbreath, A.K., et al. (2020). Nested-association mapping (NAM)-based genetic dissection uncovers candidate genes for seed and pod weights in peanut (*Arachis hypogaea*). Plant Biotechnol. J. *18*, 1457–1471. https://doi.org/10.1111/pbi.13311.

4. Li, N., Xu, R., and Li, Y. (2019). Molecular networks of seed size control in plants. Annu. Rev. Plant Biol. *70*, 435–463. https://doi.org/10.1146/annurev-arplant-050718-095851.

5. Wybouw, B., and De Rybel, B. (2019). Cytokinin-A developing story. Trends Plant Sci. *24*, 177–185. https://doi.org/10.1016/j.tplants.2018.10.012.

6. Ashikari, M., Sakakibara, H., Lin, S., Yamamoto, T., Takashi, T., Nishimura, A., Angeles, E.R., Qian, Q., Kitano, H., and Matsuoka, M. (2005). Cytokinin oxidase regulates rice grain production. Science *309*, 741–745. https://doi.org/10.1126/science.1113373.

7. Zhang, L., Zhao, Y.L., Gao, L.F., Zhao, G.Y., Zhou, R.H., Zhang, B.S., and Jia, J.Z. (2012). *TaCKX6-D1*, the ortholog of rice *OsCKX2*, is associated with grain weight in hexaploid wheat. New Phytol. *195*, 574–584. https://doi.org/10.1111/j.1469-8137.2012.04194.x.

8. Zhao, J., Bai, W., Zeng, Q., Song, S., Zhang, M., Li, X., Hou, L., Xiao, Y., Luo, M., Li, D., et al. (2015). Moderately enhancing cytokinin level by down-regulation of *GhCKX* expression in cotton concurrently increases fiber and seed yield. Mol. Breed. *35*, 60. https://doi.org/10.1007/s11032-015-0232-6.

9. Ishimaru, K., Hirotsu, N., Madoka, Y., Murakami, N., Hara, N., Onodera, H., Kashiwagi, T., Ujiie, K., Shimizu, B.I., Onishi, A., et al. (2013). Loss of function of the IAA-glucose hydrolase gene *TGW6* enhances rice grain weight and increases yield. Nat. Genet. *45*, 707–711. https://doi.org/10.1038/ng.2612.

10. Hu, Z., Lu, S.J., Wang, M.J., He, H., Sun, L., Wang, H., Liu, X.H., Jiang, L., Sun, J.L., Xin, X., et al. (2018). A novel QTL *qTGW3* encodes the GSK3/SHAGGY-like kinase OsGSK5/OsSK41 that interacts with OsARF4 to negatively regulate grain size and weight in rice. Mol. Plant *11*, 736–749. https://doi.org/10.1016/j.molp.2018.03.005.

11. Zhu, J.K. (2009). Active DNA demethylation mediated by DNA glycosylases. Annu. Rev. Genet. *43*, 143–166. https://doi.org/10.1146/annurev-genet-102108-134205.

12. Law, J.A., and Jacobsen, S.E. (2010). Establishing, maintaining and modifying DNA methylation patterns in plants and animals. Nat. Rev. Genet. *11*, 204–220. https://doi.org/10.1038/nrg2719.

13. Song, Q.X., Lu, X., Li, Q.T., Chen, H., Hu, X.Y., Ma, B., Zhang, W.K., Chen, S.Y., and Zhang, J.S. (2013). Genome-wide analysis of DNA methylation in soybean. Mol. Plant *6*, 1961–1974. https://doi.org/10.1093/mp/sst123.

14. Xing, M.Q., Zhang, Y.J., Zhou, S.R., Hu, W.Y., Wu, X.T., Ye, Y.J., Wu, X.X., Xiao, Y.P., Li, X., and Xue, H.W. (2015). Global analysis reveals the crucial roles of DNA methylation during rice seed development. Plant Physiol. *168*, 1417–1432. https://doi.org/10.1104/pp.15.00414.

15. Wang, P., Xia, H., Zhang, Y., Zhao, S., Zhao, C., Hou, L., Li, C., Li, A., Ma, C., and Wang, X. (2015). Genome-wide high-resolution mapping of DNA methylation identifies epigenetic variation across embryo and endosperm in Maize (*Zea may*). BMC Genom. *16*, 21. https://doi.org/10.1186/s12864-014-1204-7.

16. Zhang, M., Cui, G., Bai, X., Ye, Z., Zhang, S., Xie, K., Sun, F., Zhang, C., and Xi, Y. (2021). Regulatory network of preharvest sprouting resistance revealed by integrative analysis of mRNA, noncoding RNA, and DNA methylation in wheat. J. Agric. Food Chem. *69*, 4018–4035. https://doi.org/10.1021/acs.jafc.1c00050.

17. Huang, H., Liu, R., Niu, Q., Tang, K., Zhang, B., Zhang, H., Chen, K., Zhu, J.K., and Lang, Z. (2019). Global increase in DNA methylation during orange fruit development and ripening. Proc. Natl. Acad. Sci. USA *116*, 1430–1436. https://doi.org/10.1073/pnas.1815441116.

18. Pu, H., Shan, S., Wang, Z., Duan, W., Tian, J., Zhang, L., Li, J., Song, H., and Xu, X. (2020). Dynamic changes of DNA methylation induced by heat treatment were involved in ethylene signal transmission and delayed the postharvest ripening of tomato fruit. J. Agric. Food Chem. *68*, 8976–8986. https://doi.org/10.1021/acs.jafc.0c02971.

19. Bertioli, D.J., Jenkins, J., Clevenger, J., Dudchenko, O., Gao, D., Seijo, G., Leal-Bertioli, S.C.M., Ren, L., Farmer, A.D., Pandey, M.K., et al. (2019). The genome sequence of segmental allotetraploid peanut *Arachis hypogaea*. Nat. Genet. *51*, 877–884. https://doi.org/10.1038/s41588-019-0405-z.

20. Zhou, X., Han, X., Lyu, S.C., Bunning, B., Kost, L., Chang, I., Cao, S., Sampath, V., and Nadeau, K.C. (2021). Targeted DNA methylation profiling reveals epigenetic signatures in peanut allergy. JCI Insight *6*, e143058. https://doi.org/10.1172/jci.insight.143058.

21. Zhang, X., Pandey, M.K., Wang, J., Zhao, K., Ma, X., Li, Z., Zhao, K., Gong, F., Guo, B., Varshney, R.K., and Yin, D. (2021). Chromatin spatial organization of wild type and mutant peanuts reveals high-resolution genomic architecture and interaction alterations. Genome Biol. *22*, 315. https://doi.org/10.1186/s13059-021-02520-x.

22. He, L., Huang, H., Bradai, M., Zhao, C., You, Y., Ma, J., Zhao, L., Lozano-Durán, R., and Zhu, J.K. (2022). DNA methylation-free Arabidopsis reveals crucial roles of DNA methylation in regulating gene expression and development. Nat. Commun. *13*, 1335. https://doi.org/10.1038/s41467-022-28940-2.

23. Erdmann, R.M., and Picard, C.L. (2020). RNA-directed DNA methylation. PLoS Genet. *16*, e1009034. https://doi.org/10.1371/journal.pgen.1009034.

24. Hayashi, K.I., Arai, K., Aoi, Y., Tanaka, Y., Hira, H., Guo, R., Hu, Y., Ge, C., Zhao, Y., Kasahara, H., and Fukui, K. (2021). The main oxidative inactivation pathway of the plant hormone auxin. Nat. Commun. *12*, 6752. https://doi.org/10.1038/s41467-021-27020-1.

25. Clevenger, J., Chu, Y., Scheffler, B., and Ozias-Akins, P. (2016). A developmental transcriptome map for allotetraploid *Arachis hypogaea*. Front. Plant Sci. *7*, 1446. https://doi.org/10.3389/fpls.2016.01446.

26. Wang, P., Shi, S., Ma, J., Song, H., Zhang, Y., Gao, C., Zhao, C., Zhao, S., Hou, L., Lopez-Baltazar, J., et al. (2018). Global methylome and gene expression analysis during early peanut pod development. BMC Plant Biol. *18*, 352. https://doi.org/10.1186/s12870-018-1546-4.

27. Lin, J.Y., Le, B.H., Chen, M., Henry, K.F., Hur, J., Hsieh, T.F., Chen, P.Y., Pelletier, J.M., Pellegrini, M., Fischer, R.L., et al. (2017). Similarity between soybean and Arabidopsis seed methylomes and loss of non-CG methylation does not affect seed development. Proc. Natl. Acad. Sci. USA *114*, E9730–E9739. https://doi.org/10.1073/pnas.1716758114.

28. Zhong, S., Fei, Z., Chen, Y.R., Zheng, Y., Huang, M., Vrebalov, J., McQuinn, R., Gapper, N., Liu, B., Xiang, J., et al. (2013). Single-base resolution methylomes of tomato fruit development reveal epigenome modifications associated with ripening. Nat. Biotechnol. *31*, 154–159. https://doi.org/10.1038/nbt.2462.

29. Zhang, M., Kimatu, J.N., Xu, K., and Liu, B. (2010). DNA cytosine methylation in plant development. J. Genet. Genom. *37*, 1–12. https://doi.org/10.1016/S1673-8527(09)60020-5.

30. Lang, Z., Wang, Y., Tang, K., Tang, D., Datsenka, T., Cheng, J., Zhang, Y., Handa, A.K., and Zhu, J.K. (2017). Critical roles of DNA demethylation in the activation of ripening-induced genes and inhibition of ripening-repressed genes in tomato fruit. Proc. Natl. Acad. Sci. USA *114*, E4511–E4519. https://doi.org/10.1073/pnas.1705233114.

31. Wang, L., Zheng, K., Zeng, L., Xu, D., Zhu, T., Yin, Y., Zhan, H., Wu, Y., and Yang, D.L. (2022). Reinforcement of CHH methylation through RNA-directed DNA methylation ensures sexual reproduction in rice. Plant Physiol. *188*, 1189–1209. https://doi.org/10.1093/plphys/kiab531.

32. Kawakatsu, T., Nery, J.R., Castanon, R., and Ecker, J.R. (2017). Dynamic DNA methylation reconfiguration during seed development and germination. Genome Biol. *18*, 171. https://doi.org/10.1186/s13059-017-1251-x.

33. Zhou, Z., Liu, C., Qin, M., Li, W., Hou, J., Shi, X., Dai, Z., Yao, W., Tian, B., Lei, Z., et al. (2022). Promoter DNA hypermethylation of *TaGli-γ-2.1* positively regulates gluten strength in bread wheat. J. Adv. Res. *36*, 163–173. https://doi.org/10.1016/j.jare.2021.06.021.

34. Wang, H., Zhu, Y., Yuan, P., Song, S., Dong, T., Chen, P., Duan, Z., Jiang, L., Lu, L., and Duan, H. (2021). Response of wheat DREB transcription factor to osmotic stress based on DNA methylation. Int. J. Mol. Sci. *22*, 7670. https://doi.org/10.3390/ijms22147670.

35. Lei, M., Zhang, H., Julian, R., Tang, K., Xie, S., and Zhu, J.K. (2015). Regulatory link between DNA methylation and active demethylation in Arabidopsis. Proc. Natl. Acad. Sci. USA *112*, 3553–3557. https://doi.org/10.1073/pnas.1502279112.

36. Hirose, N., Takei, K., Kuroha, T., Kamada-Nobusada, T., Hayashi, H., and Sakakibara, H. (2008). Regulation of cytokinin biosynthesis, compartmentalization and translocation. J. Exp. Bot. *59*, 75–83. https://doi.org/10.1093/jxb/erm157.

37. Brugière, N., Humbert, S., Rizzo, N., Bohn, J., and Habben, J.E. (2008). A member of the maize isopentenyl transferase gene family, Zea mays *isopentenyl transferase 2* (*ZmIPT2*), encodes a cytokinin biosynthetic enzyme expressed during kernel development. Plant Mol. Biol. *67*, 215–229. https://doi.org/10.1007/s11103-008-9312-x.

38. Ljung, K. (2013). Auxin metabolism and homeostasis during plant development. Development *140*, 943–950. https://doi.org/10.1242/dev.086363.

39. Korasick, D.A., Enders, T.A., and Strader, L.C. (2013). Auxin biosynthesis and storage forms. J. Exp. Bot. *64*, 2541–2555. https://doi.org/10.1093/jxb/ert080.

40. Bernardi, J., Lanubile, A., Li, Q.B., Kumar, D., Kladnik, A., Cook, S.D., Ross, J.J., Marocco, A., and Chourey, P.S. (2012). Impaired auxin biosynthesis in the defective endosperm18 mutant is due to mutational loss of expression in the *ZmYuc1* gene encoding endosperm-specific YUCCA1 protein in maize. Plant Physiol. *160*, 1318–1328. https://doi.org/10.1104/pp.112.204743.

41. Mashiguchi, K., Tanaka, K., Sakai, T., Sugawara, S., Kawaide, H., Natsume, M., Hanada, A., Yaeno, T., Shirasu, K., Yao, H., et al. (2011). The main auxin biosynthesis pathway in Arabidopsis. Proc. Natl. Acad. Sci. USA *108*, 18512–18517. https://doi.org/10.1073/pnas.1108434108.

42. Fenn, M.A., and Giovannoni, J.J. (2021). Phytohormones in fruit development and maturation. Plant J. *105*, 446–458. https://doi.org/10.1111/tpj.15112.

43. Zhu, L., Jiang, B., Zhu, J., and Xiao, G. (2022). Auxin promotes fiber elongation by enhancing gibberellic acid biosynthesis in cotton. Plant Biotechnol. J. *20*, 423–425. https://doi.org/10.1111/pbi.13771.

44. Dello Ioio, R., Nakamura, K., Moubayidin, L., Perilli, S., Taniguchi, M., Morita, M.T., Aoyama, T., Costantino, P., and Sabatini, S. (2008). A genetic framework for the control of cell division and differentiation in the root meristem. Science *322*, 1380–1384. https://doi.org/10.1126/science.1164147.

45. Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170.

46. Krueger, F., and Andrews, S.R. (2011). Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. Bioinformatics *27*, 1571–1572. https://doi.org/10.1093/bioinformatics/btr167.

47. Wu, H., Xu, T., Feng, H., Chen, L., Li, B., Yao, B., Qin, Z., Jin, P., and Conneely, K.N. (2015). Detection of differentially methylated regions from whole-genome bisulfite sequencing data without replicates. Nucleic Acids Res. *43*, e141. https://doi.org/10.1093/nar/gkv715.

48. Young, M.D., Wakefield, M.J., Smyth, G.K., and Oshlack, A. (2010). Gene ontology analysis for RNA-seq: accounting for selection bias. Genome Biol. *11*, R14. https://doi.org/10.1186/gb-2010-11-2-r14.

49. Xie, C., Mao, X., Huang, J., Ding, Y., Wu, J., Dong, S., Kong, L., Gao, G., Li, C.Y., and Wei, L. (2011). Kobas 2.0: a web server for annotation and identification of enriched pathways and diseases. Nucleic Acids Res. *39*, W316–W322. Web Server issue. https://doi.org/10.1093/nar/gkr483.

50. Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. *14*, R36. https://doi.org/10.1186/gb-2013-14-4-r36.

51. Pertea, M., Pertea, G.M., Antonescu, C.M., Chang, T.C., Mendell, J.T., and Salzberg, S.L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat. Biotechnol. *33*, 290–295. https://doi.org/10.1038/nbt.3122.

52. Li, B., and Dewey, C.N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinf. *12*, 323. https://doi.org/10.1186/1471-2105-12-323.

53. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550. https://doi.org/10.1186/s13059-014-0550-8.

54. Li, Z., Zhang, X., Zhao, K., Zhao, K., Qu, C., Gao, G., Gong, F., Ma, X., and Yin, D. (2021). Comprehensive transcriptome analyses reveal candidate genes for variation in seed size/weight during peanut (Arachis hypogaea L.) domestication. Front. Plant Sci. *12*, 666483. https://doi.org/10.3389/fpls.2021.666483.

55. Dong, N.Q., Sun, Y., Guo, T., Shi, C.L., Zhang, Y.M., Kan, Y., Xiang, Y.H., Zhang, H., Yang, Y.B., Li, Y.C., et al. (2020). UDP-glucosyltransferase regulates grain size and abiotic stress tolerance associated with metabolic flux redirection in rice. Nat. Commun. *11*, 2629. https://doi.org/10.1038/s41467-020-16403-5.

56. Chen, Y., Su, D., Li, J., Ying, S., Deng, H., He, X., Zhu, Y., Li, Y., Chen, Y., Pirrello, J., et al. (2020). Overexpression of bHLH95, a basic helix-loop-helix transcription factor family member, impacts trichome formation via regulating gibberellin biosynthesis in tomato. J. Exp. Bot. *71*, 3450–3462. https://doi.org/10.1093/jxb/eraa114.

57. Lister, R., Mukamel, E.A., Nery, J.R., Urich, M., Puddifoot, C.A., Johnson, N.D., Lucero, J., Huang, Y., Dwork, A.J., Schultz, M.D., et al. (2013). Global epigenomic reconfiguration during mammalian brain development. Science *341*, 1237905. https://doi.org/10.1126/science.1237905.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
RNA sequence data for H1314 S1 (B)	This study	NCBI Sequence Read Archive: SRR24119575
RNA sequence data for H1314 S1 (A)	This study	NCBI Sequence Read Archive: SRR24119576
RNA sequence data for H1314 S4 (A)	This study	NCBI Sequence Read Archive: SRR24119588
RNA sequence data for H1314 S4 (B)	This study	NCBI Sequence Read Archive: SRR24119587
RNA sequence data for H2014 S1 (A)	This study	NCBI Sequence Read Archive: SRR24119586
RNA sequence data for H2014 S1 (B)	This study	NCBI Sequence Read Archive: SRR24119585
RNA sequence data for H2014 S4 (A)	This study	NCBI Sequence Read Archive: SRR24119584
RNA sequence data for H2014 S4 (B)	This study	NCBI Sequence Read Archive: SRR24119583
Whole-genome bisulphite sequencing data for H1314 S1 (A)	This study	NCBI Sequence Read Archive: SRR24119590
Whole-genome bisulphite sequencing data for H1314 S1 (B)	This study	NCBI Sequence Read Archive: SRR24119589
Whole-genome bisulphite sequencing data for H1314 S4 (A)	This study	NCBI Sequence Read Archive: SRR24119582
Whole-genome bisulphite sequencing data for H1314 S4 (B)	This study	NCBI Sequence Read Archive: SRR24119581
Whole-genome bisulphite sequencing data for H2014 S1 (A)	This study	NCBI Sequence Read Archive: SRR24119580
Whole-genome bisulphite sequencing data for H2014 S1 (B)	This study	NCBI Sequence Read Archive: SRR24119579
Whole-genome bisulphite sequencing data for H2014 S4 (A)	This study	NCBI Sequence Read Archive: SRR24119578
Whole-genome bisulphite sequencing data for H2014 S4 (B)	This study	NCBI Sequence Read Archive: SRR24119577
cultivated peanut cv. Tifrunner reference genome	Bertioli et al.[19]	https://v1.legumefederation.org/data/v1/Arachis_hypogaea/Tifrunner.gnm1.KYV3/arahy.Tifrunner.gnm1.KYV3.genome_main.fna.gz
Experimental models: Organisms/strains		
H2014	Zhang et al.[21]	N/A
H1314	Zhang et al.[21]	N/A
Software and algorithms		
Trimmomatic	Bolger et al.[45]	https://github.com/usadellab/Trimmomatic
Bismark	Krueger and Andrews.[46]	https://github.com/FelixKrueger/Bismark
DSS	Wu et al.[47]	https://bioconductor.org/packages/release/bioc/html/DSS.html
GOseq	Young et al.[48]	https://bioconductor.org/packages/release/bioc/html/goseq.html
KOBAS	Xie et al.[49]	http://kobas.cbi.pku.edu.cn/
Tophat2	Kim et al.[50]	http://ccb.jhu.edu/software/tophat/index.shtml
StringTie	Pertea et al.[51]	https://ccb.jhu.edu/software/stringtie/

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
RSEM	Li and Dewey.[52]	https://deweylab.github.io/RSEM/
DESeq2	Love et al.[53]	https://bioconductor.org/packages/release/bioc/html/DESeq2.html

RESOURCE AVAILABILITY

Lead contact

Further information should be directed to the lead contact, Dongmei Yin (yindm@henau.edu.cn).

Materials availability

This study did not generate new unique materials.

Data and code availability

- The Illumina sequencing raw data (RNA-seq and whole-genome bisulphite sequencing) generated in this study have been deposited in NCBI (National Center for Biotechnology Information) under SRA (Sequence Read Archive: SRR24119575-SRR24119576, SRR24119583-SRR24119588 for seed transcriptomic data from H1314 and H2014; SRR24119577-SRR24119582, SRR24119589-SRR24119590 for seed methylation data from H1314 and H2014), and are publicly available as of the date of publication. In addition, the cultivated peanut cv. Tifrunner reference genome was used in this study.[19] Accession numbers are listed in the key resources table.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Plant materials

Line H1314 (mutant) was isolated from an ethyl- methanesulphonate (EMS)-mutagenised M_2 population of peanut accession H2014 (wild type, WT), which is an elite peanut cultivar with medium-size seeds, and is used for genetic analysis following successive reproduction.[21] Both peanut lines were provided by the Peanut Genetics & Breeding Lab of Henan Agricultural University, and were grown at the experimental station of Zhengzhou, China, from May to September 2019–2021. Normal field management was conducted during peanut growth and development. Peanut seed development from zygotic embryo (peanut gynophore or peg entering into soil) to seed maturation always undergoes three phases, morphogenesis, maturation and mature, corresponding to cell division, cell expansion, and loss of water, respectively. Our previous study has showed that peanut seeds have maximum size, protein and oil accumvulation at ~60 days after pegging (DAP, S4), which is approximately at seed late maturation stage, and have continual and extensive cell division at 15 DAP (S1).[54] So fresh subterranean pods at 15 and 60 DAP, corresponding to stage 1 (S1) and 4 (S4) of seed development, respectively, were harvested from 5–10 plants, and washed thoroughly with sterilized deionized water. Then enough seeds were quickly collected, and stored at -80°C for subsequent research.

METHOD DETAILS

Endogenous phytohormone determination

Approximately 0.3 g (fresh weight) seed samples at S1 and S4 stages from mutant and wild type plants were weighted, ground into powder in liquid nitrogen for auxin, CK and GA determination. Extraction and quantification of auxin (26 compounds including IAA, IBA and ICA) and CKs (36 compounds including IP, tZ, cZ and DZ) were conducted as described by Dong et al.,[55] while GAs levels (GA_1, GA_3, GA_{4-9}, GA_{15}, GA_{19}, GA_{20}, GA_{24}, GA_{29}, GA_{34}, GA_{51} and GA_{53}) were measured following the description of Chen et al.[56] Three biological replicates were included for each stage or sample. These analyses were performed by Wuhan Metware Biotechnology Co., Ltd. (Wuhan, China) using the AB Sciex QTRAP 6500 liquid chromatography-tandem mass spectrometry (LC-MS/MS) platform.

Cytological observation

To measure cell size, fresh seeds at S4 were collected from both H1314 and H2014, fixed in FAA (50% ethanol, 5% glacial acetic acid, 5% formaldehyde) for 48 h at room temperature, then dehydrated in an ethanol series (30%, 50%, 75%, 85%, 90%, 95%, 100% ethanol, 1−2 h per stage). After treatment with xylene, the seed samples were embedded in wax and sliced into 5 μm thin sections with a pathology slicer (Leica, RM2016, Germany). Dewaxing treatment of seed tissue sections was performed in xylene twice for 20 min each time, 100% ethanol twice for 5 min each time, 75% ethanol for 5 min, followed by washing with tap water. After staining with hematoxylin and eosin dyesolution (BIOPPLE, G1003) for 3−5 min, the sections were observed under a light microscope (Nikon Eclipse E100, Nikon, Japan). Photos were collected and analyzed using the software CaseViewer2.4 and Image-Pro Plus 6.0 (Media cybernetics, Inc., USA).

Whole-genome bisulphite sequencing data

Firstly, genomic DNA was isolated from seed samples of peanut accession H2014 and the H1314 mutant at S1 and S4 using a DNeasy Plant Maxi Kit (Qiagen, Germany). DNA integrity, purity and concentration were determined by agarose gel electrophoresis, NanoPhotometer spectrophotometer (IMPLEN, CA, USA) and a Qubit DNA Assay Kit (Life Technologies, CA, USA), respectively. After DNA fragmentation and purification, we performed end modification and adenylation, adapter ligation, bisulphite conversion, fragment selection, PCR amplification, and quality control. The resulting libraries with an insert size of 200−400 bp were sequenced at Novogene Co., Ltd. (Beijing, China) on an Illumina Novaseq 6000 sequencer, and 150 bp paired-end reads were generated.

Secondly, raw data were pre-processed using Trimmomatic software[45] by removing low-quality reads (q <15), and clean reads were mapped to the reference genome of the allotetraploid cultivated peanut (*A. hypogaea* cv. Tifrunner) using the Bismark software.[46] We identified methylated cytosines using the bismark_methylation_extractor tool. Briefly, duplicated reads resulting from PCR amplification were removed, and methylation was considered to have occurred when cytosines of bisulphite-treated reads were aligned to the reference genome and had ≥ 5×coverage and q-value ≤0.05 in sequencing libraries. For each methylated cytosine, methylation level (ML) represents the number of reads with cytosine methylation divided by total reads covering the corresponding cytosine, and this was further corrected with the bisulphite non-conversion rate according to a previous study.[57]

Thirdly, Differentially methylated regions (DMRs) were identified using DSS software (v2.12.0) based on the Beta-Binomial distribution.[47] Briefly, the regions that have at least three differentially methylated cytosines (P <0.01, Fisher's exact test), a length of ≥50 bp, and methylation level change with ≥0.5, 0.3, and 0.2 for CG, CHG and CHH, respectively, and a Benjamin–Hochberg corrected P value < 0.05 were selected for further analysis. And then the two regions within 100-bp of each other were merged to a new one. Genes in which gene body (from TSS to TES), or promoter regions (2 kb upstream from the TSS) have an overlap of ≥1 bp with DMRs were defined to be DMR-related genes. Furthermore, Gene Ontology (GO) enrichment analysis of genes related to DMRs was implemented by the GOseq R package,[48] and terms with corrected p-value <0.05 were considered significantly enriched. Finally, statistical enrichment of Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway was tested by using KOBAS software.[49]

RNA-seq, data processing and gene annotation

Approximately, ~3 μg total RNA was extracted from each seed sample using an RNA Prep Pure Plant Kit (Tiangen Co., Beijing, China), and treated with DNaseI (Thermo Fisher Scientific Inc., Grand Island, NY). RNA concentration and integrity were assessed using a Qubit RNA Assay Kit and a Qubit 2.0 Fluorimeter (Life Technologies) and an RNA Nano 6000 Assay Kit with a Bioanalyzer 2100 system (Agilent Technologies, CA, USA), respectively. After mRNA isolation, fragmentation, synthesis of double-stranded cDNAs, end modification, fragment selection, PCR amplification, product purification and quality assessment, sequencing libraries with an insert size of ~350 bp were obtained and submitted to an Illumina Novaseq 6000 sequencer, and paired-end reads with a length of 150 bp were generated.

After strict quality control was performed by excluding reads containing adapter sequences, low-quality reads with ≥10% ambiguous nucleotides (Ns), and ≥50% of bases with a quality scores (Q-score) ≤20, large numbers of clean reads were generated for each sequenced seed sample. Meanwhile, we also assessed Q20, Q30, GC content and sequence duplication level for the clean data. The resulting clean reads were mapped to the assembly genome of the cultivated peanut cv. Tifrunner[19] using Tophat2.[50]

Furthermore, we conducted transcript assembly and quantification of gene expression abundance based on StringTie[51] and the fragments per kilobase per million reads (FPKM) method using RSEM,[52] respectively. Additionally, we carried out annotation of gene function using several databases including Nr (NCBI non-redundant protein sequences), Pfam (Protein family) and Swiss-Prot (A manually annotated and reviewed protein sequence database).

DEGs analyses, GO and KEGG enrichment

Since each of the seed samples from H1314 and H2014 included two biological replicates at S1 and S4, identification of DEGs between mutant and wild type was performed using the DESeq2 R package (1.20.0), and genes with an adjusted *p*-value < 0.05 and 2-fold change of expression level were assigned as differentially expressed.[53] Enrichment analyses of GO terms and KEGG pathways were carried out using the GOseq R package based on the Wallenius non-central hyper-geometric distribution[48] and KOBAS,[49] respectively.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analysis used in comparative genome and transcriptome analysis can be found in the relevant sections of the method details. Data for quantification analyses are presented as mean ± standard deviation (SD). Statistical analyses were performed using two-way ANOVA with Tukey's multiple comparisons test, unpaired *t*-test (for *t*-test, **, P-value <0.01. *, P-value <0.05. NS, not significantly) for comparisons between two groups. Data was analyzed and plotted using GraphPad Prism version 8 and SAS software.