

RESEARCH ARTICLE

Open Access



DNA methylation and its effects on gene expression during primary to secondary growth in poplar stems

Yang Zhang^{1†}, Cong Liu^{1†}, He Cheng¹, Shuanghui Tian¹, Yingying Liu¹, Shuang Wang¹, Huaxin Zhang², Muhammad Saqib³, Hairong Wei⁴ and Zhigang Wei^{2*}

Abstract

Background: As an important epigenetic mark, 5-methylcytosine (5mC) methylation is involved in many DNA-dependent biological processes and plays a role during development and differentiation of multicellular organisms. However, there is still a lack of knowledge about the dynamic aspects and the roles of global 5mC methylation in wood formation in tree trunks. In this study, we not only scrutinized single-base resolution methylomes of primary stems (PS), transitional stems (TS), and secondary stems (SS) of *Populus trichocarpa* using a high-throughput bisulfite sequencing technique, but also analyzed the effects of 5mC methylation on the expression of genes involved in wood formation.

Results: The overall average percentages of CG, CHG, and CHH methylation in poplar stems were ~ 53.6%, ~ 37.7%, and ~ 8.5%, respectively, and the differences of 5mC in genome-wide CG/CHG/CHH contexts among PS, TS, and SS were statistically significant ($p < 0.05$). The evident differences in CG, CHG, and CHH methylation contexts among 2 kb proximal promoters, gene bodies, and 2 kb downstream regions were observed among PS, TS, and SS. Further analysis revealed a perceptible global correlation between 5mC methylation levels of gene bodies and transcript levels but failed to reveal a correlation between 5mC methylation levels of proximal promoter regions and transcript levels. We identified 653 and 858 DMGs and 4978 and 4780 DEGs in PS vs TS and TS vs SS comparisons, respectively. Only 113 genes of 653 DMGs and 4978 DEGs, and 114 genes of 858 DMGs and 4780 DEG were common. Counterparts of some of these common genes in other species, including *Arabidopsis thaliana*, are known to be involved in secondary cell wall biosynthesis and hormone signaling. This indicates that methylation may directly modulate wood formation genes and indirectly attune hormone signaling genes, which in turn impact wood formation.

Conclusions: DNA methylation only marginally affects pathway genes or regulators involved in wood formation, suggesting that further studies of wood formation should lean towards the indirect effects of methylation. The information and data we provide here will be instrumental for understanding the roles of methylation in wood formation in tree species.

Keywords: DNA methylation, Gene expression, Primary stems, Transition stems, Secondary stems, *Populus trichocarpa*

* Correspondence: zhigangwei1973@163.com

[†]Yang Zhang and Cong Liu contributed equally to this work.

²Research Center of Saline and Alkali Land of State Forestry and Grassland Administration, Chinese Academy of Forestry, Beijing 100091, People's Republic of China

Full list of author information is available at the end of the article



© The Author(s). 2020 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Background

Wood is the most abundant biomass produced by plants, especially trees, and can serve as a renewable resource for energy, pulp, paper products, and building materials [1]. In most trees, wood originates from vascular cambium, the secondary meristem located between tree barks and woody trunks; vascular cambium produces undifferentiated xylem mother cells inwardly and bark cells outwardly. For this reason, cambium activity is the most important determining factor for wood accumulation. Present knowledge indicates that the differentiation of vascular cambium into xylem mother cells is controlled by plant hormones and HD-ZIP III transcription factors [2]. After that, xylem mother cells undergo a series of biological processes, including cell division and expansion, secondary wall formation, lignification, and finally programmed cell death, to produce secondary xylem, known as wood [3, 4]. The coordinated activation of secondary wall biosynthesis in xylem mother cells to produce wood is mediated by a transcriptional network composed of secondary wall NAC and MYB master switches and their downstream transcription factors [1, 2, 5–8]. However, the contribution of epigenetic regulation during this process is still unclear.

DNA methylation, a key epigenetic modification, typically involves the addition of a methyl group to the fifth carbon of cytosine to produce 5-methylcytosine (5mC) in eukaryotic genomes [9, 10]. Although the relationship between DNA methylation and its effect on gene expression is complex [11, 12], an increasing body of evidence suggests that DNA methylation plays a role in various biological processes during plant growth and development [12, 13], such as morphogenesis [14], gender determination [15, 16], vegetative propagation [17], and response to abiotic stress [18–21]. In plants, cytosine methylation is primarily found in three sequence contexts: CG, CHG, and CHH (where H = A, T or C) [22]. Moreover, DNA methylation exhibits tissue specific patterns in plants. For example, in *Arabidopsis thaliana*, about 6% of cytosines are methylated in immature floral tissues [23], while 24% CG, 6.7% CHG, and 1.7% CHH are methylated in young plants [24]. In rice, whole genome methylation patterns are similar among mature leaves, embryos, and seedling shoots and roots, but hypomethylation levels are correlated with expression levels of genes that are preferentially expressed in endosperm [25]. Patterns of 5mC in long terminal repeat (LTR) transposable elements differ between rice leaves and roots [26] and affect neighboring gene expression in *A. thaliana* [27, 28]. Tissue-specific characteristics of genome methylation are also evident in natural populations of Chinese white poplar [29]. Although DNA methylation is purported to play an important role in wood formation [30, 31], the mechanisms by which

DNA methylation alter the expression of xylogenetic genes have not been elucidated. Moreover, tissue-specific methylation patterns in the transitional zones between vascular cambium and secondary wood have not been characterized.

The transition from primary to secondary growth can be easily observed in the stems of less than one-year old poplar trees with multiple developmental stages. For instance, stems near apical meristems are generally soft and green due to the presence of a multitude of cells with primary cell walls; in contrast, stems in basal portions are stiff and woody owing to the presence of a large fraction of secondary xylem cells that have undergone cell wall thickening and lignification. Stems in the middle are in a transitional stage between primary and secondary growth. For this reason, vertical segments of developing stems from less than one-year old trees constitute an ideal experimental system for investigating epigenetic regulatory mechanisms of wood formation [32, 33]. To date, no focused study of DNA methylation and its effects on gene expression in different developmental stages of stems has been conducted in tree species. In this study, we generated high coverage genome-wide maps of cytosine methylation at single-nucleotide resolution and transcriptomic profiles of *Populus trichocarpa* stems in various developmental stages varying from predominantly primary to secondary growth. This study was designed to collect data and gain insight into four problems: (i) the genomic landscape of the different developmental stem methylomes; (ii) the changes in the methylomes associated with different stem developmental stages; (iii) an evaluation of relationships between methylome changes and expression of wood formation genes; (iv) the identification of wood formation genes that are subjected to epigenetic regulation. The epigenetic and RNA-seq data acquired constitute valuable genetic resources, and the results and conclusions drawn from the data and analysis will be instrumental for further studies of both the epigenetic and molecular regulatory mechanisms of wood formation.

Results

Morphological and histochemical changes in *P. trichocarpa* stems

To verify the rationality of the classification of the main stems of poplar into different developmental stages using the plastochron indices method, we determined the developmental stages of internodes two (IN2), four (IN4), and eight (IN8) using histochemical staining. Toluidine blue-O and phloroglucinol-HCl were used to stain lignin while calcofluor white was used to stain cellulose in xylem vessel elements. Because the vascular bundles in IN2 comprised mainly of primary xylem and phloem tissues that were formed from procambial cells, toluidine

blue-O and phloroglucinol-HCl staining in the cross sections of IN2 were nearly undetectable (Fig. 1b and c), and calcofluor white staining in IN2 sections was also weak (Fig. 1d). In IN4, the secondary vascular cambium has emerged and produced secondary walls. As a result, the lignin stained by either toluidine blue-O or phloroglucinol-HCl was clearly discernible (Fig. 1e and f), and the cellulose stained by calcofluor white was also more obvious (Fig. 1g) than in IN2 (Fig. 1d). In the stem segments of IN8, the secondary xylem had increasingly accumulated, phloem fibers had emerged, and both were lignified. As a result, the intensities of toluidine blue-O, phloroglucinol-HCl, and calcofluor white staining in the cross sections of IN8 (Fig. 1h-j) were much more striking than in IN2 or IN4 (Fig. 1b-g). Therefore, IN2, IN4, and IN8, representing the stages of primary stems (PS), transitional stems (TS), and secondary stems (SS) from

primary growth to secondary growth, respectively, were used for further analysis. To avoid getting into a state of uncertainty by virtue of using multiple cross-section tissues, we harvested only the primary xylem upon peeling tree bark and focused our studies on DNA methylation and genomic aspects of xylogenesis.

The expression levels of genes involved in DNA methylation and demethylation in *P. trichocarpa* stems

To determine whether variations in DNA methylation exist among PS, TS, and SS, we first used qRT-PCR to globally scrutinize the expression levels of genes involved in DNA methylation. We focused on the DNA methylation genes *PtrMET1A/B*, *PtrDRM1/2-A-C*, *PtrCMT3-A-C*, and *PtrDDM1-A/B* and the DNA demethylation genes *PtrDME-A/B*, *PtrDEMETER-LIKE 2-A/B*, and *PtrROS1*. As shown in Fig. 2, the expression levels of *PtrMET1-B*,

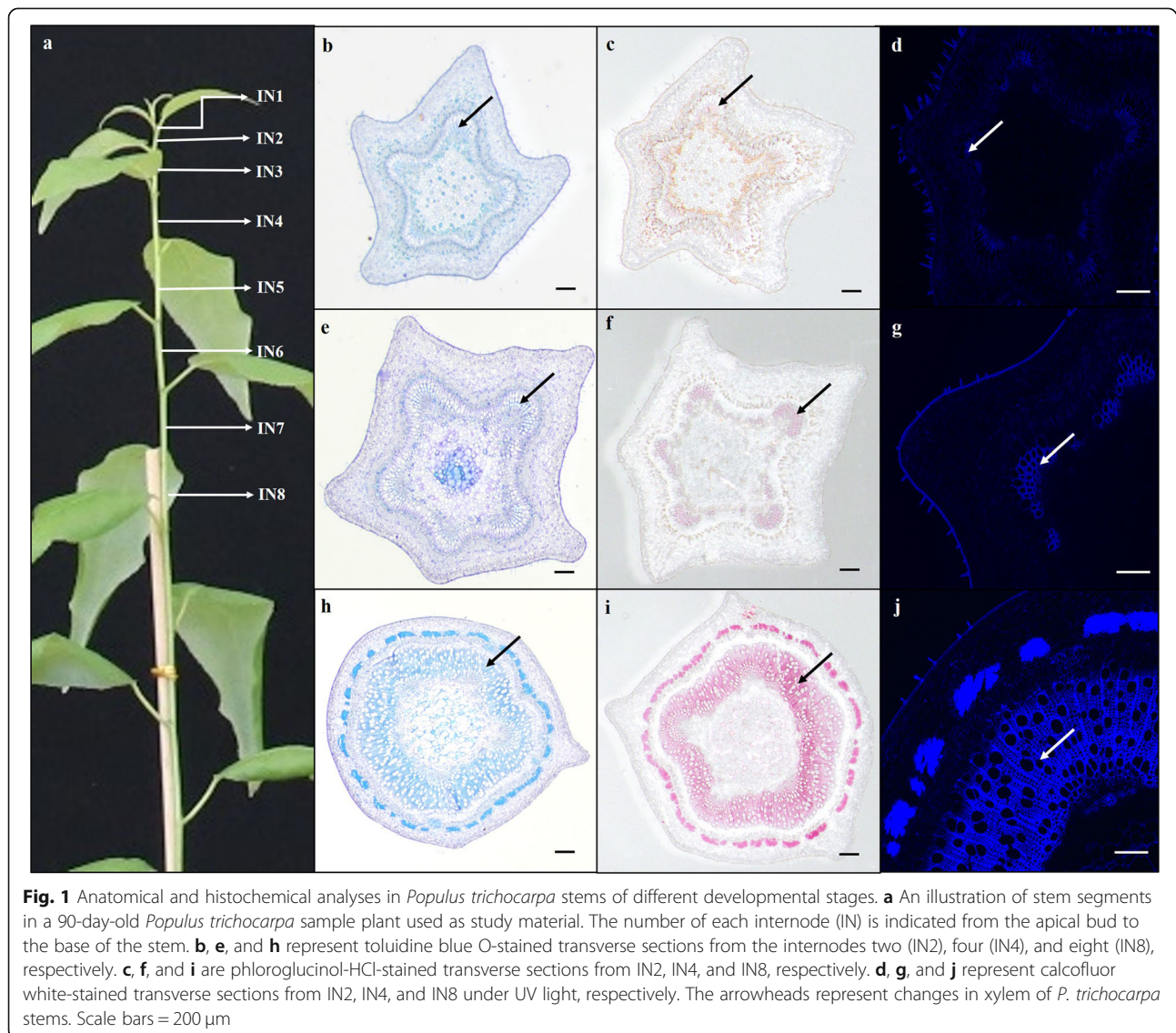
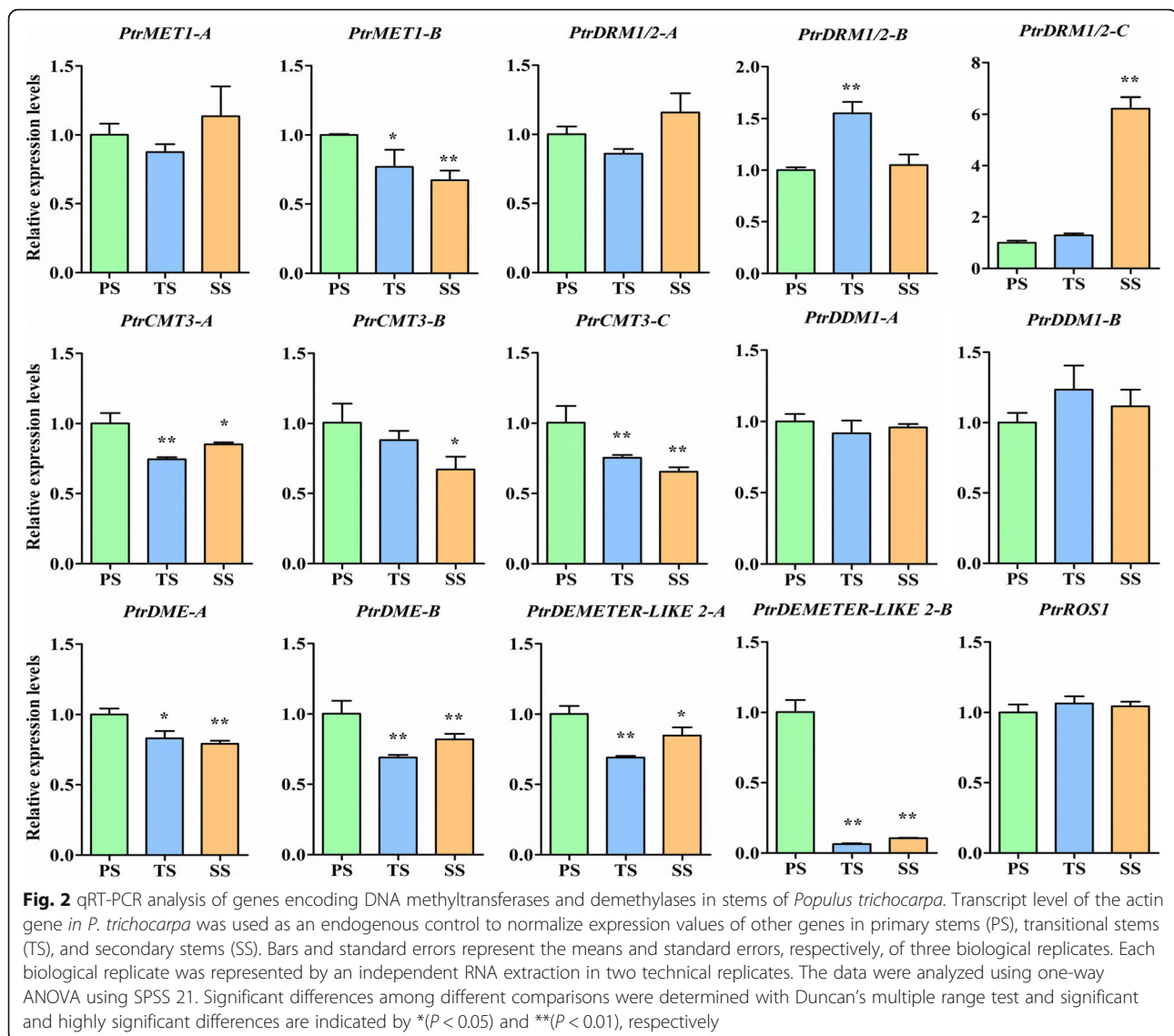


Fig. 1 Anatomical and histochemical analyses in *Populus trichocarpa* stems of different developmental stages. **a** An illustration of stem segments in a 90-day-old *Populus trichocarpa* sample plant used as study material. The number of each internode (IN) is indicated from the apical bud to the base of the stem. **b**, **e**, and **h** represent toluidine blue O-stained transverse sections from the internodes two (IN2), four (IN4), and eight (IN8), respectively. **c**, **f**, and **i** are phloroglucinol-HCl-stained transverse sections from IN2, IN4, and IN8, respectively. **d**, **g**, and **j** represent calcofluor white-stained transverse sections from IN2, IN4, and IN8 under UV light, respectively. The arrowheads represent changes in xylem of *P. trichocarpa* stems. Scale bars = 200 μ m



PtrCMT3-A, and *PtrCMT3-C* were significantly different among PS, TS, and SS. *PtrDRM1/2-C* had significantly higher and *PtrCMT3-B* had significantly lower expression levels in SS than in both PS and TS. The expression level of *PtrDRM1/2-B* in TS was significantly higher than in PS and SS. However, there were no statistically significant differences in the expression levels of *PtrMET1-A*, *PtrDRM1/2-A*, *PtrDDM1-A*, and *PtrDDM1-B* among PS, TS, and SS. Of the expressed genes involved in DNA demethylation, only *PtrDME-A* and *PtrDEMETER-LIKE 2-A* exhibited significant differences in expression levels among PS, TS, and SS. The expression levels of *PtrDME-B* and *PtrDEMETER-LIKE 2-B* in TS and SS exhibited significant differences compared to PS. However, the expression levels of these three demethylation genes had no obvious differences between TS and SS. Moreover, there were no significant differences in the expression

levels of *PtrROS1* among PS, TS, and SS. In summary, the differential expression of these genes across three developmental stages suggests that genomic DNA methylation patterns may be altered during the wood formation process.

Whole-Genome Bisulfite Sequencing (WGBS) of the *P. trichocarpa* genome

Variations in the expression levels of genes involved in DNA methylation suggest that genomic DNA methylation levels might be different across PS, TS, and SS. To investigate the genomics methylation levels of poplar in the stems of different developmental stages, we performed bisulfite sequencing of genomic DNA extracted from PS, TS, and SS using the Illumina HiSeq 2500 platform. We then decoded and analyzed the corresponding methylomes. A total of

99.5–115.6 million raw sequencing paired reads were obtained for each biological replicate (Table 1), covering the whole genome of *P. trichocarpa* with a depth between 30.35–34.29-fold. Raw reads were then subjected to a series of filtering criteria to ensure data quality, and 99.82–99.88% of the reads were retained for further analysis. The reads from each sample were mapped to the *P. trichocarpa* reference genome with a mapping rate of 75.74–80.27% (Table 1).

DNA methylation landscapes of *P. trichocarpa* genome

Conversion rates were calculated by aligning reads to the unmethylated lambda DNA added to the total DNA before applying bisulfite treatment. Conversion rates of genomic DNA of PS, TS, and SS were on average 99.51, 99.52, and 99.50%, respectively, these rates were used to conduct binomial tests to exclude those 5mCs that may be the result of non-conversion of cytosines in our bisulfite treatment or sequencing errors resulting from the base calling process. Then, we obtained on average 14,773,999, 16,392,099, and 16,852,157 mCs for the PS, TS, and SS genomes, respectively (Additional file 1). The PS genome harbored ~ 11.96, 47.64, 28.85, and 4.85% methylated C at the total sequenced C, CG, CHG, and CHH sites, respectively. Likewise, the TS genome contained ~ 13.33, 49.89, 31.53, and 5.88% methylated C while the SS genome contained ~ 13.55, 48.80, 30.80, and 6.46% methylated C, respectively, at the total sequenced C, CG, CHG, and CHH sites (Additional file 1). We also found that, regardless of developmental stage (PS, TS, or SS), ~ 45% of CG and ~ 65% of CHG sites were lowly methylated (0–10%) while ~ 40% of CG and ~ 24% of CHG sites were highly methylated (90–100%) (Fig. 3a and b); in contrast, ~ 83% of CHH sites were lowly methylated (0–10%) (Fig. 3c) and less than 1% of CHH sites were highly methylated (90–100%). These results suggest that nearly half of CG methylation sites are either hypomethylated or hypermethylated, nearly two thirds of CHG methylation sites are hypermethylated, and the majority of CHH sites are hypomethylated in poplar stems.

As an important methylation characteristics of a genome, the proportions of mCG, mCHG, and mCHH on total mC sites have species and tissues specificity. Thus, we not only identified the distribution patterns of mC sites in mCG, mCHG, and mCHH contexts among PS, TS, and SS, but we also compared the mC site distribution patterns of poplar stems with *A. thaliana* [23, 24, 34], rice [35], and apple [36]. The overall distribution patterns of mC sites in mCG, mCHG, and mCHH were illustrated using Chromosome 1 in the PS genome (Fig. 3d). The distribution of mCs in the other 18 chromosomes were also determined on sense and antisense strands (Additional files 2, 3, and 4).

We found that PS, TS, and SS exhibited nearly same distribution patterns of total mC sites in three methylation contexts as those in apple on the whole. However, the mCs exhibited different distribution patterns in *P. trichocarpa* compared to *A. thaliana* and rice, especially in mCG and mCHH contexts (Fig. 3e). In PS, 5mC was found more frequently at CHH sites (43.87%) than at CG (28.42%) or CHG (27.71%) sites. In TS and SS, CHH methylation rates increased to 48.47 and 51.93%, respectively, indicating that the CHH methylation rate increases in accordance with the progression of secondary growth and development. Accordingly, the CG methylation rates in TS and SS decreased to 25.95 and 24.22%, respectively, while the CHG methylation rates decreased to 25.61 and 23.87%, respectively (Additional file 5), suggesting that the levels of these two methylation contexts negatively correlate with the progression of secondary growth. The comparisons of PS, TS, and SS among CG, CHG, and CHH methylation rates revealed that there were significant differences in mCG, mCHG, and mCHH contexts among PS, TS, and SS (Additional file 6).

Regardless of PS, TS, and SS, the poplar genome showed a relatively lower methylation level within gene-rich regions compared to a relatively high degree of methylation within transposable element (TE)-rich regions (Fig. 4). Moreover, the gene-rich regions with few or no TEs exhibited a relatively less methylation levels

Table 1 Description of the bisulfite sequencing (BS-Seq) data of early developing stems (3-month-old) in *Populus trichocarpa*

Sample		Raw reads	Clean reads	Sequence Depth	Mapped Reads	Mapped Rate (%)
PS	Rep1	105,344,898	105,190,862 (99.85%)	30.79	80,968,657	76.97
	Rep2	99,553,450	99,425,308 (99.87%)	30.35	79,809,380	80.27
	Rep3	107,558,146	107,425,046 (99.88%)	32.16	84,583,644	78.74
TS	Rep1	107,666,760	10,749,4930 (99.84%)	31.51	82,870,707	77.09
	Rep2	109,561,376	109,409,538 (99.86%)	31.51	82,864,569	75.74
	Rep3	115,622,622	115,474,758 (99.87%)	34.14	89,790,043	77.76
SS	Rep1	113,318,946	113,174,330 (99.87%)	34.29	90,186,749	79.69
	Rep2	107,441,448	107,246,874 (99.82%)	32.09	84,393,945	78.69
	Rep3	115,564,860	115,357,460 (99.82%)	33.95	89,294,118	77.41

Note: primary stems (PS), transitional stems (TS), and secondary stems (SS)



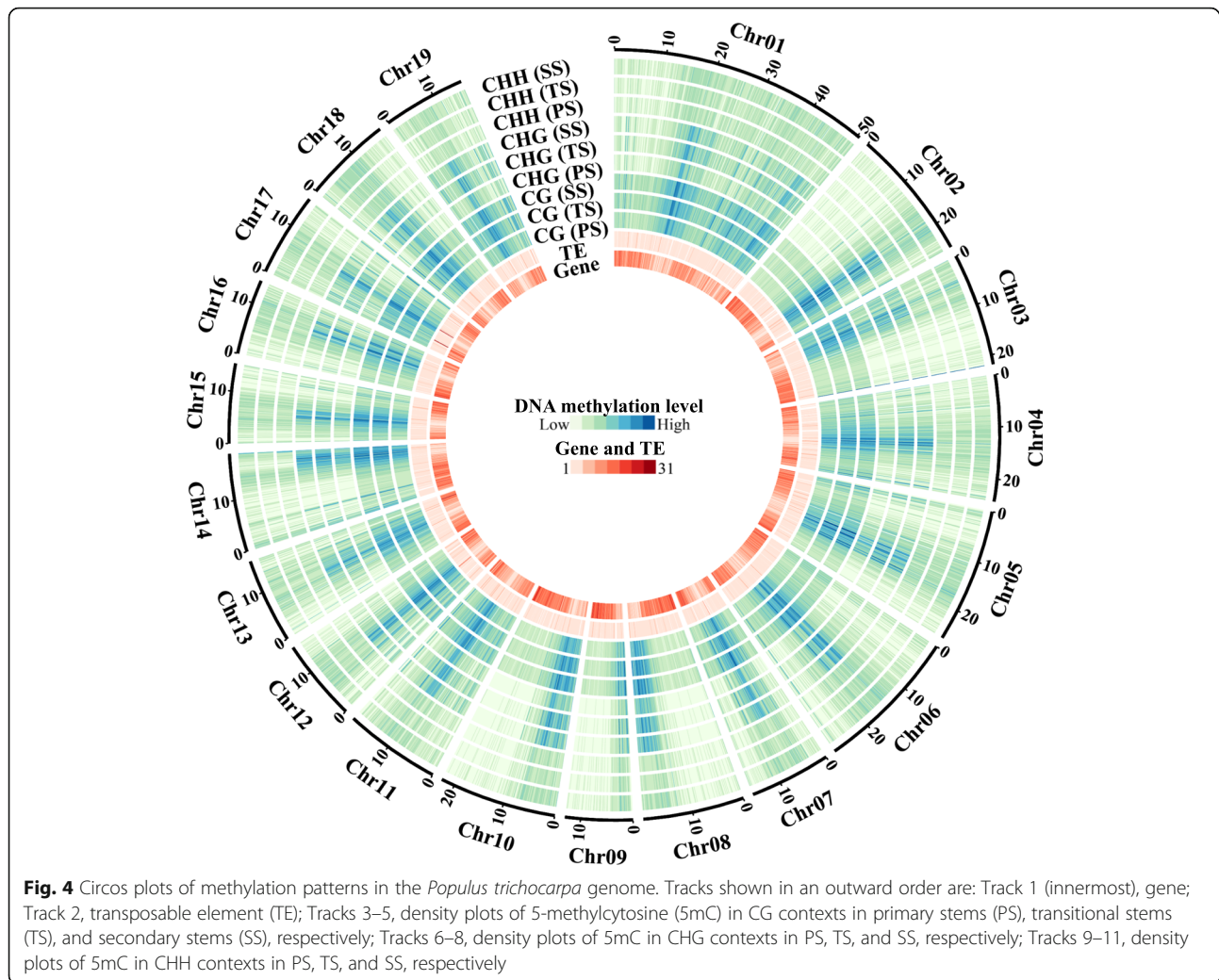
(Fig. 4) as compared to the gene-rich regions with more TEs. We did not find large-scale differences in the genomes of *P. trichocarpa* stems from different stages.

Genomic methylation patterns in *P. trichocarpa* stems

Given the existence of tissue level variation in DNA methylation in the *P. trichocarpa* genome [30], we further explored the methylation profiles of PS, TS, and SS within different genomic regions; this included different genic and intergenic regions, especially repetitive regions containing various transposable elements (TEs) such as long terminal repeats (LTR), long interspersed nuclear elements (LINE), short interspersed nuclear elements (SINE), and DNA transposons (DNA). In PS, TS, and SS, CG and CHG methylation levels were higher than CHH methylation levels in each of the genomic regions mentioned above (Fig. 5a and b). There were significant differences in methylation in the CG and CHG contexts when various specific genomic regions were compared. For example, PS/TS/SS_Gene body of CG verse PS/TS/SS_Gene body of CHG methylation (Fig. 5b). In addition, methylation levels in CG, CHG, and CHH contexts were slightly higher in TS and SS than in PS (Fig. 5a and b). LTRs had the highest methylation levels in all three methylation contexts (CG, CHG, and CHH) in PS, TS, and SS (Fig. 5a). In contrast, SINEs had the lowest methylation levels in all three contexts and stages of stem development. LINEs had modest methylation in all three contexts of methylation and also

three developmental stages. Further research found that LTR Gypsy, LTR Caulimovirus, LINE L1, and DNA CMC-EnSpm, the predominant type of transposable element sequence in *P. trichocarpa* genomes [37], had higher methylation levels than others in the stems of *P. trichocarpa* (Additional file 7). In addition, the LTR Copia and LTR Gypsy super families had no distinct differences in their methylation levels across PS, SS, and TS, which resembles their relatively invariant methylation levels across seven tissues (vegetative bud, male inflorescence, female catkin, leaf, root, xylem, and phloem) of *P. trichocarpa* as observed earlier [30]. Among different genic regions, the 5'UTR and 3'UTR had much lower methylation levels than other regions; promoters and 2 kb downstream regions had higher methylation levels than other regions in all three methylation contexts in PS, TS, and SS (Fig. 5b).

We also found that methylation levels changed during stem development in *P. trichocarpa*. In TEs and genic regions, methylation levels of CG, CHG, and CHH contexts were increased in TS and SS compared to PS (Fig. 5c and d), and methylation levels in CG and CHG contexts were highest in TS. However, the methylation levels in CHH contexts were highest in SS. As for TE regions, the CG context had the highest methylation level, and CHH had the lowest level of the three methylation contexts in PS, TS, and SS (Fig. 5c). In both CG and CHG contexts, TEs had higher methylation levels than 2 kb upstream and 2 kb downstream regions. However,



there were no conspicuous differences in CHH methylation levels among TEs, 2 kb upstream, and 2 kb downstream regions in all three tissues. Additional studies showed that several TE super families, including LINE L1, DNA CMC-EnSpm, DNA hAT-Tag1, and DNA hAT-Tlp100, had higher CG and CHG methylation levels in TEs than their 2 kb upstream and 2 kb downstream regions (Additional file 8). In addition, the LTR Copia and LTR Gypsy super families had no distinct differences in three CG, CHG, and CHH methylation levels among TEs and their 2 kb upstream and downstream regions in all three tissues.

As shown in Fig. 5d, the three methylation contexts were ranked consistently from highest to lowest methylation as CG, CHG, and CHH, no matter which genic regions or tissue types were considered. Moreover, promoter regions had higher methylation levels compared with either gene bodies or 2 kb downstream regions in all three stem tissues. It was notable that gene bodies had lower methylation levels, especially for CHG and

CHH, as compared with either the promoter regions or the 2 kb downstream regions. To compare methylation levels in the three genomic contexts in different genic regions across multiple tissues, multiple comparison testing was conducted; significant differences among difference comparisons are provided in Additional file 9. Within the promoter regions, there were no significant differences in CG methylation levels among PS, TS, and SS (Additional file 9A). There were significant differences in CHG methylation levels between PS and TS and also between PS and SS (Additional file 9B) and significant differences in CHH methylation levels among PS, TS, and SS (Additional file 9C). Within the gene bodies, there were significant differences in CG contexts between TS and SS (Additional file 9D), in CHG contexts between PS and TS and between PS and SS (Additional file 9E), and in CHH contexts among PS, TS, and SS (Additional file 9F). Within the 2 kb downstream regions, there were significant differences in CG methylation levels between PS and SS, in CHG between

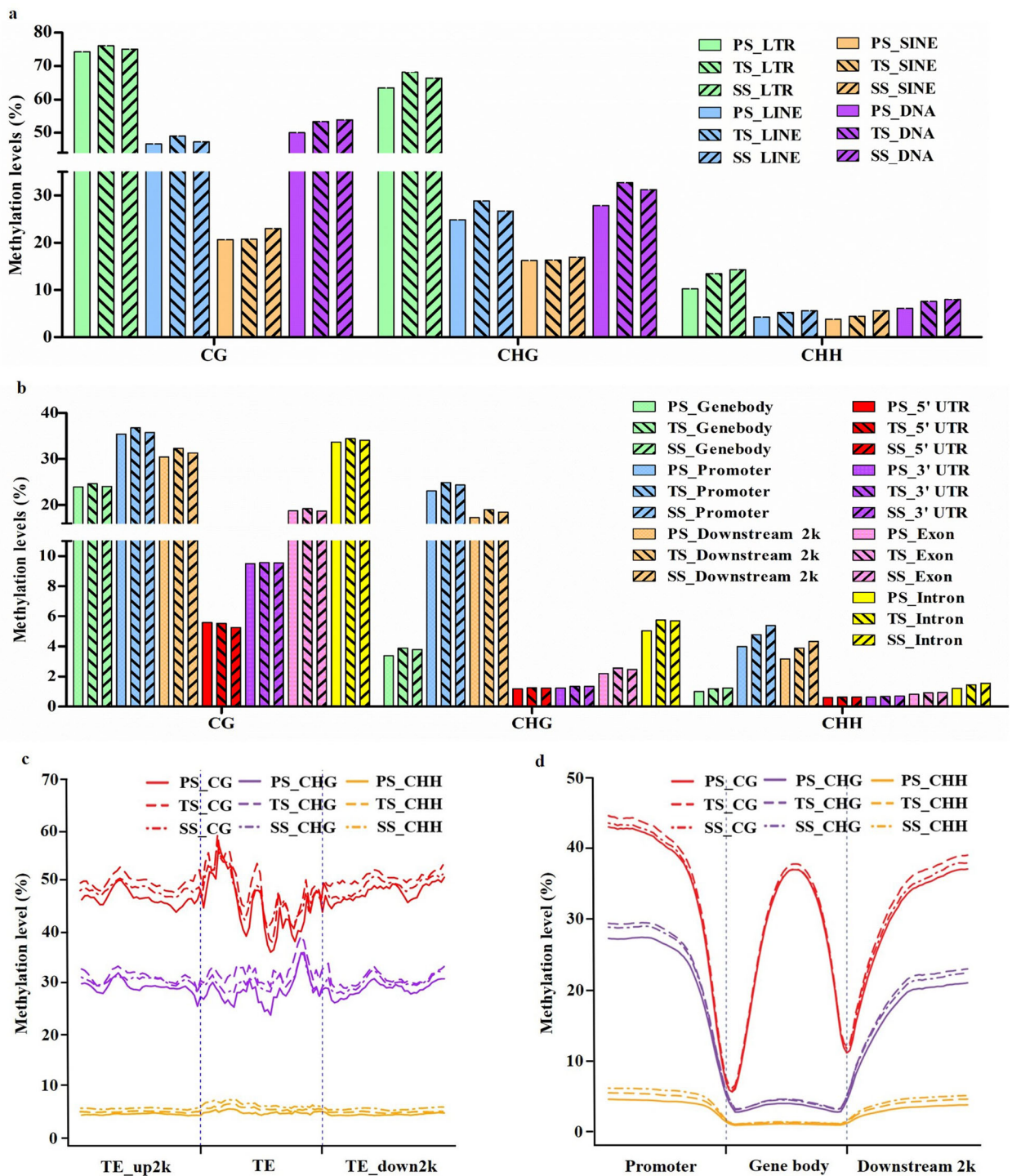
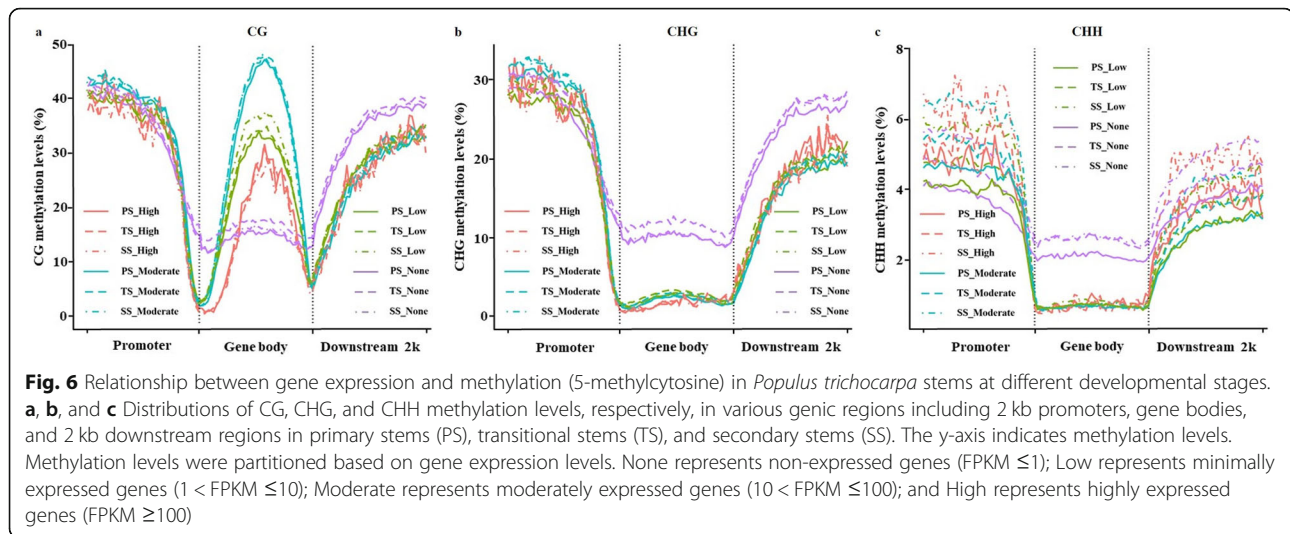


Fig. 5 DNA methylation patterns in different genic regions. **a** Methylation levels in different types of transposable elements (TEs), including long terminal repeats (LTR), long interspersed nuclear elements (LINE), short interspersed nuclear elements (SINE), and DNA transposons (DNA) in primary stems (PS), transitional stems (TS), and secondary stems (SS). **b** Methylation levels of gene features, including promoter, 2 kb downstream region, and gene body regions with 5'UTR, exon, intron, and 3'UTR regions in stems of poplar. **c** Methylation levels among TE regions and their 2 kb upstream and downstream regions in stems of poplar. **d** Distribution of methylation levels among gene features, including promoter, gene body, and 2 kb downstream regions in stems of poplar. The y-axis indicates methylation levels



PS and TS and between PS and SS, and in CHH methylation levels between any combination of PS, TS, and SS (Additional file 9G-I). However, despite being statistically significant, the variations observed in methylation levels are limited.

Relation between DNA methylation and gene expression

To investigate the potential influence of gene methylation on gene expression during wood formation, transcriptome profiling of PS, TS, and SS was conducted using the same materials used for methylome analysis. Based on the FPKM gene expression levels, we classified all genes into non-expressed or expressed genes; the latter were further divided into three groups based on their expression level: low, moderate, or high. We then scrutinized the differences in CG, CHG, and CHH methylation levels within the proximal promoters, gene bodies, and 2 kb downstream regions of genes having different expression levels. As shown in Fig. 6, the four differentially expressed groups had different methylation levels in the three methylation contexts among proximal promoters, gene bodies, and 2 kb downstream regions. As expected, non-expressed genes had the highest CHG and CHH methylation levels in gene body regions and the highest CG, CHG, and CHH methylation levels within 2 kb downstream regions in PS, TS, and SS. In contrast, non-expressed genes had the lowest CHH methylation levels and moderate CG and CHG methylation levels in upstream 2 kb promoter regions in PS, TS, and SS. Interestingly, regardless of their expression levels, expressed genes had higher CG, but lower CHG and CHH, methylation levels than non-expressed genes in gene bodies in PS, TS, and SS. Expressed genes with moderate expression levels had the highest CG methylation levels within gene bodies, and the highest CG and CHG methylation levels in promoter regions in PS, TS,

and SS. These results suggest that genes with different expression levels correspond to different CG, CHG, and CHH methylation levels in different genic regions. Overall, CG and CHG methylation patterns in expressed genes are similar in promoter regions while CHG and CHH methylation patterns in expressed genes are similar in gene bodies and downstream regions.

To further evaluate the relationships between genic methylation and gene expression, genes with transcriptomic profiles were classified into unmethylated (None) and methylated groups; genes in the methylated group were further divided into three subgroups: the bottom third were referred to as the low-methylation subgroup (Low), the middle third as the moderate-methylation subgroup (Moderate), and the top third as the high-methylation group (High) based on their methylation levels. The results suggest that methylations in gene bodies have more obvious correlation on gene expression than those in promoters and 2 kb downstream regions in all three methylation contexts (Additional file 10). The proportion of genes with the lowest methylation levels within promoters and with the highest methylation levels within 2 kb downstream regions were the lowest among the three methylation levels, no matter which type of methylation contexts they were. Moreover, the greatest proportion of genes were moderately methylated in gene bodies in all three methylation contexts. These results suggest that methylation level and context have different effects on gene expression depending on their genic location.

To further study the relationship between genic methylation and gene expression, Spearman correlation analysis was performed between methylation levels in the whole gene frame (bodies ± 2 kb flanking regions) and gene expression levels. As shown in Additional file 11, the overall correlation rho was low regardless of the methylation types and the developmental stages. However, the rho can

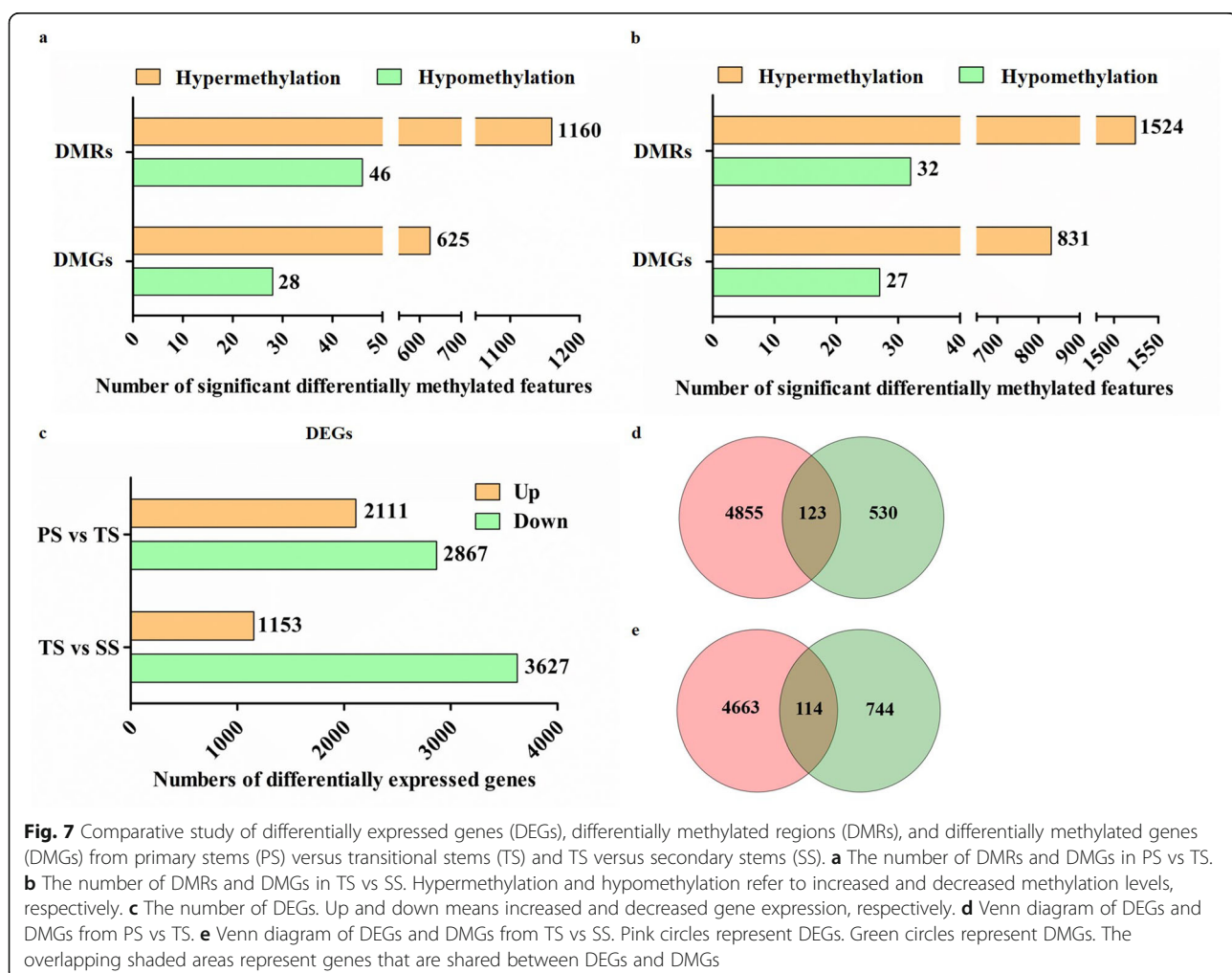
reach 0.25 in gene bodies for CG methylation in PS, TS, or SS. Such a rho may indicate that methylation levels in a small fraction of gene bodies have relatively higher correlation with their expression levels.

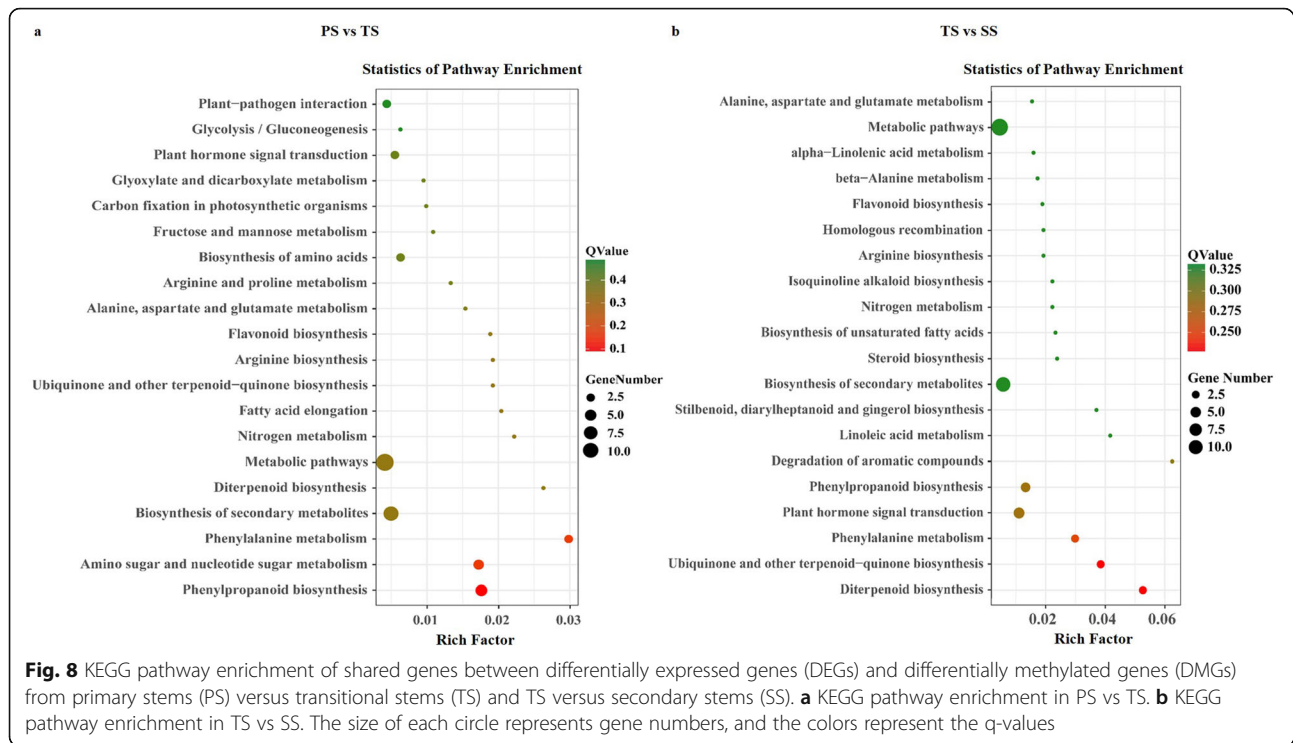
Widespread dynamic gene methylation in *P. trichocarpa* stems

To determine the relationship between developmental stages and methylation, we analyzed the differentially methylated regions (DMRs) between PS and TS and between TS and SS. For PS vs TS, the 1206 DMRs (1160 hypermethylated and 46 hypomethylated regions) (Fig. 7a) overlapped 653 differentially methylated genes (DMGs) (625 hypermethylated and 28 hypomethylated genes). For TS vs SS, the 1556 DMRs (1524 hypermethylated and 32 hypomethylated regions) overlapped the 858 DMGs (831 hypermethylated and 27 hypomethylated genes) (Fig. 7b). At the same time, we also identified 4978 differentially expressed genes (DEGs) from PS vs TS and 4780 DEGs from TS vs SS (Fig. 7c). However, there are only 123 common genes between the 653

DMGs and the 4978 DEGs from PS vs TS, and 114 common genes between the 858 DMGs and the 4780 DEGs from TS vs SS (Fig. 7d and e).

Functional enrichment analysis on the 123 common genes revealed the enrichment of phenylpropanoid biosynthesis, phenylalanine metabolism, and the biosynthesis of secondary metabolites (Fig. 8a) from PS vs TS. KEGG pathway enrichment analyses on the rest of the DMGs from PS vs TS revealed the enrichment of pathways including pyrimidine metabolism, isoflavonoid biosynthesis, caffeine metabolism, and homologous recombination (Additional file 12A). Enrichment analyses with the DEGs from PS vs TS revealed pathways for phenylpropanoid biosynthesis, phenylalanine metabolism, and the biosynthesis of secondary metabolites (Additional file 12B), suggesting that DNA methylation participates in the initial stage of secondary cell wall formation. The same analysis was applied to the 114 genes common to DMGs and DEGs from TS vs SS, and the results showed that diterpenoid biosynthesis, ubiquinone and other terpenoid-quinone biosynthesis,





phenylalanine metabolism, plant hormone signal transduction, and phenylpropanoid biosynthesis pathways were among those enriched (Fig. 8b). It was surprising that the enriched pathways in the rest of the DMGs from TS vs SS only consisted of brassinosteroid biosynthesis and caffeine metabolism pathways (Additional file 12C); the remaining DEGs from TS vs SS revealed multiple metabolic pathways, including biosynthesis of secondary metabolites, phenylpropanoid biosynthesis, and flavonoid biosynthesis (Additional file 12D), again indicating that DNA methylation contributes to secondary wood formation and metabolites during the TS to SS transition.

Since transcription factors (TFs) lie at the center of gene regulation, we identified TFs with pronounced alternations in methylation levels. We found 32 TFs in the DMGs from PS vs TS and 39 TFs in the DMGs from TS vs SS which shared 11 TFs with the 229 DEGs from PS vs TS and 6 TFs with the 245 DEGs from TS vs SS (Additional files 13 and 14). For example, we found the TFs *NAC056*, *MYB52*, *NTL9*, *WRKY27*, and *MYB106* from PS vs TS (Additional file 13) and the TFs *bZIP21*, *MYB52*, *ANAC047*, and *HRD* from TS vs SS (Additional file 14).

Discussion

Whole-genome bisulfite sequencing (WGBS) is a powerful technology for studying genome-wide DNA methylation patterns at single-nucleotide resolution. To date, an increasing number of plant species have been studied using WGBS, ranging from model species like *A.*

thaliana [23, 24, 34] to some agriculturally important crops like rice [38, 39], maize [40], soybean [41], tomato [42], spruce [43], oil palm [44], and *P. trichocarpa* [45]. However, there is still a lack of understanding of the roles of methylation in secondary growth and development of tree species. Here, we decoded the single-base poplar stem methylome by WGBS and obtained an overall average of ~48.78%, ~30.39%, and ~5.72% methylation levels in CG, CHG, and CHH contexts (Additional file 1), respectively. Although methylation levels in poplar stems fell into the ranges described for other angiosperms (CG: ~30.5% to ~92.5%, CHG: ~9.3 to ~81.2%, and CHH: ~1.1% to ~18.8%) [46], they were on the lower end in each methylation context. The relatively low methylation levels in poplar stems maybe due, in part, to the fact that the *P. trichocarpa* genome (~422.9 Mb) is small compared to that of other plant species; previous studies have demonstrated that DNA methylation level is positively correlated with genome size [43, 46]. For example, apple, which has a moderately sized genome (~742 Mb), had ~53.6%, ~37.7%, and ~8.5% methylation levels in CG, CHG, and CHH contexts, respectively [47]. In *A. thaliana*, CG sites show a bimodal distribution and tend to be either unmethylated or highly methylated, whereas CHG and CHH sites are rarely methylated at high levels [23, 24]. This phenomenon implies that distinct mechanisms are responsible for maintaining different methylation types. As reported, CG methylation is copied faithfully during DNA replication, whereas CHG and CHH methylation are perpetually targeted by

histone methylation and/or noncoding RNAs [13]. Interestingly, in *P. trichocarpa*, we found bimodal distribution patterns (unmethylated or methylated) for both CG and CHG methylation (Fig. 3a and b), suggesting that CHG methylation is maintained more robustly and copied more faithfully in poplar than in *A. thaliana*. Additionally, our results showed that methylated sites were concentrated on non-CG sites in the genome of *P. trichocarpa*, particularly on CHH sites, which is similar to observations in other trees [36, 48]. We also found that methylation levels of poplar stems were negatively correlated with gene numbers (Fig. 4), which was consistent with previous studies [49]. Furthermore, CG, CHG, and CHH methylation levels were positively correlated with TE density (Fig. 4) in the similar ways as what were demonstrated as in apple [47]. Some types of TEs, including LTR Copia and LTR Gypsy, are primarily distributed in pericentromeric regions of chromosomes [50]; we found that methylation of poplar stems also peaked in the centromere and pericentromeric regions as demonstrated in Fig. 3d and Additional files 2, 3, and 4. This finding has also been reported in other plant species [46, 47], indicating that hypermethylation may contribute to the maintenance of chromosomal stability and segregation.

Tissue differentiation coupled with distinct methylation levels is a universal phenomenon in all multicellular organisms. For example, methylation levels differ among seven distinct tissue types in *P. trichocarpa* [30], including vegetative bud, male inflorescence, female catkins, leaves, roots, xylem, and phloem, and in leaves [51], buds [52], and xylem [53] in natural populations of Chinese white poplar. Such a differentiation in methylation is essential for plant growth and development [54]. In this study, we found that the percentage of minimally methylated (0–10%) CG, CHG, and CHH sites decreased in TS and SS than PS (Fig. 3a, b and c). Therefore, mCG, mCHG, and mCHH sites could be established more stably in TS and SS. We also found that CG and CHH methylation levels showed significant differences among PS, TS, and SS. For CHG methylation, although PS showed more noticeable differences when compared with either TS or SS, there were no significant differences between TS and SS (Additional file 6). The methylation levels in CG contexts were highest in TS, and the methylation levels in CHH contexts were highest in SS (Fig. 5). These results suggest that, to some degree, genome methylation is in accord with wood formation in poplar stems and undergoes an obvious alteration at the initiation of secondary growth. It is well known that demethylases plays an important role in DNA methylation. In TS and SS, the reduced expression of demethylases (*PtrDME-A*, *PtrDME-B*, *PtrDEMETER-LIKE 2-A*, and *PtrDEMETER-LIKE 2-B*) might contribute to the

increased methylation level (Fig. 2). Moreover, the enzymatic activities of two proteins, DRM1 and DRM2, are responsible for asymmetric CHH methylation [55]. *PtrDRM1/2-B* and *PtrDRM1/2-C* had significant higher expression levels in TS and SS (Fig. 2), respectively, than other developmental stages. The increased CHH methylation level may have something to do with the increased expression of *PtrDRM1/2* in TS and SS.

It has been reported that DNA methylation can repress gene expression [56]. However, the relationship between DNA methylation and gene transcription is more nuanced than initially realized. For example, rice promoter methylation repressed gene expression only in some heavily methylated gene loci; on the contrary, gene body methylation was positively, rather than negatively, correlated with gene expression [35]. In this study, we found that the relationship between DNA methylation and gene expression was complicated by genic regions, methylation contexts, and developmental states. Previous results showed that methylation tends to suppress gene transcription, and non-methylated genes had higher expression levels [31]. However, in our study, we found that the expressed genes in PS, TS, and SS had higher CG but lower CHG and CHH methylation levels in gene bodies than non-expressed genes (Fig. 6). In addition, Methylation in gene bodies corresponded more strongly with gene expression levels than methylation in either promoter or 2 kb downstream regions. In general, we found that, compared to expressed genes, non-expressed genes in PS, TS, and SS had the lowest CHH methylation in promoter regions, the highest CHG and CHH methylation levels in gene bodies, and the highest CG and CHG methylation levels in 2 kb downstream regions (Fig. 6). DNA methylation in promoters has been shown to affect gene expression levels, and the influence of promoter methylation on gene expression increases as the methylation level escalates [57–59]. In this study, we found that, regardless of developmental stage (PS, TS, and SS), CHH methylation levels in promoter regions positively corresponded to gene expression levels (Fig. 6c) and thus may be used as an epigenetic mark for gene expression in poplar stems. However, CHH methylation levels in the promoter regions of active genes were not consistently higher than those of other gene groups, particularly genes with moderate expression levels. Therefore, caution needs to be taken when CHH methylation is used as an epigenetic mark of gene expression.

We identified only 123 common genes in 653 DMGs and 4978 DEGs from PS vs TS, and 114 common genes in 858 DMGs and 4780 DEGs from TS vs SS (Fig. 7d and e), which suggests that developmental processes are coupled with 5mC to some degree. The counterparts of some common genes in *A. thaliana* are involved in cell growth (*CCR4* [60], *MAP 70-1* [61], and *UGT85A* [62]), secondary

cell wall biosynthesis (*NAC056* [63], *PAL1* [64], *FLA11* [65], *C4H* [66], and *MYB52* [67]), and hormone and signal transduction (*WRKY27* [68], *UGT84B1* [69], *anac047* [70], *PIP1* [71], and *CYP716A1* [72]) (Additional files 15 and 16), indicating that methylation can affect some individual pathway or regulatory genes involved in wood formation. For example, expression of *PtrPAL2* (Potri.008G038200) increased dramatically with the change of DNA methylation sites in PS vs TS (Additional file 15). As we also know, the deamination of phenylalanine by *PAL* is the first step in monolignol biosynthesis [64]. Some *PAL* genes are specifically expressed in differentiating xylem, while others are expressed in many other tissues. Homologs of *PtrPAL2* in *Populus fremontii* × *angustifolia* [73] and in *Populus* ESTs [74, 75] also suggests that it is xylem specific. *PtrCAH1* (Potri.013G157900) was also observed with the similar changes in TS vs SS (Additional file 16). *C4H* converts cinnamic acid into 4-hydroxycinnamic acid, a precursor for many phenylpropanoids including flavonoids, phytoalexins, and monolignols [76]. *PtrCAH1* transcripts are abundant in differentiating xylem, suggesting that it is important in monolignol biosynthesis [66]. In addition, our functional enrichment analysis showed that DNA methylation participated in the initial stage of secondary cell wall formation, but thereafter DNA methylation mainly influenced secondary metabolites. This was further verified by analyzing methylated TFs. Of the 32 and 39 methylated TFs identified from PS vs TS and TS vs SS, respectively, 11 are among the 229 DEGs from PS vs TS and 6 are among 245 DEGs from TS vs SS (Additional files 13 and 14). However, among these 16 methylated TFs, there were three *PtrMYB52* genes (Potri.008G089700, Potri.012G039400, and Potri.015G033600) whose counterparts in *A. thaliana* have been shown to regulate secondary cell wall formation [67]. This supports the idea that DNA methylation may affect wood formation by modulating transcription networks during the transition from primary growth to wood formation. According to an earlier study, differences in the expression of specific genes with unique methylation patterns, rather than relative methylation levels between the two tissue types, plays a critical role in wood biosynthesis [31]. The location of cytosine methylation within *MYB* and *NAC* genes might differentially affect the abundances of their transcripts, which is applicable to *PtrMYB52* gene too. Through integrated analysis of genomic, DNA methylomic, and transcriptomic differences between cultivated and wild rice, primary DNA sequence divergence has been shown to be the major determinant of methylational differences at the whole genome level, but DNA methylational differences alone can only account for limited gene expression variation between cultivated and wild rice [35]. Although we failed to detect large scale methylation of wood formation genes during the transition from primary to secondary growth, we still observed the role of methylation in wood

formation. It is possible that methylation may also affect wood formation indirectly by modulating hormone and signaling transduction; as shown (Fig. 8), the common genes between DMGs and DEGs from a comparison of TS vs SS were enriched in hormone signal transduction pathways. There is evidence that DNA methylation can affect the signaling processes of various hormones, including salicylic acid (SA) under biotic stress [77] and auxin, abscisic acid (ABA), SA and ethylene under abiotic stress [47]. Changes in methylation were observed particularly in the bodies of expressed genes and to a lesser extent in transposable elements. Together, DEGs and DMRs were significantly enriched in genes related to phytohormone metabolism or signaling pathways [19], suggesting that an indirect influence of methylation on wood formation may exist.

Conclusion

Our study provided DNA methylomes from multiple poplar stem tissues varying from predominantly primary to secondary growth, and then characterized the correlation between DNA methylation with gene expression during *P. trichocarpa* stem development. Our results indicate that DNA methylation only marginally affects pathway genes and regulators involved in wood formation, suggesting that further studies of wood formation should lean towards studying the indirect effects of methylation. The information and data provided here will be instrumental for understanding the roles of methylation in wood formation in tree species.

Methods

Plant materials

A few plantlets of *P. trichocarpa* clone Nisqually-1, whose genome was sequenced [78] early, were obtained from the Shanghai Institute for Biological Sciences, Chinese Academy of Sciences, and vegetatively propagated in our lab using tissue culture [79]. The plantlets were planted in humus soil and grown under 16 h/8 h day/night photoperiod at 25 °C in the greenhouse at Northeast Forestry University for 90 days. Then, 63 trees were used as experimental materials for anatomical and histological analysis, transcriptomic profiling, and DNA methylation sequencing. The lengths of all internodes (IN) from the apical buds to the bases of the main stems of all trees were measured. As illustrated in Fig. 1a, the IN2, IN4, and IN8 internodes, representing the PS, TS, and SS stages from primary growth to secondary growth, respectively, were used for anatomical and histological analysis. Other internodes were immediately frozen in liquid nitrogen and stored at -80 °C for qRT-PCR, RNA-seq, and DNA methylation sequencing. Three independent biological replicates were used for each of the above experiments.

Anatomical and histological analysis of poplar stems

Approximately 3-mm-long segments of each internode sample from PS, TS, and SS were fixed in FAA buffer (50% ethanol, 5% acetic acid and 3.7% formaldehyde) and then embedded in paraffin. Transverse sections (8- μ m thick) were cut from the embedded tissues with a sliding microtome (HM340E, Microm, Germany). Microscopic sections were stained with 0.025% (m/v) toluidine blue-O for 30 s, washed with dH₂O, mounted on glass slides, and then examined using bright field microscopy (BX43, Olympus, Japan). The lignin present in the secondary walls in microsections was detected by histological staining with 2.5% phloroglucinol in 12 N HCl for 3 min and then observed using a microscope (Stereo Lumar.V12, Zeiss, Germany). Cellulose was stained with 0.25 μ g/mL calcofluor white for 2 min and then visualized and photographed with a UV fluorescence microscope.

DNA methylation analysis using qRT-PCR

Previous studies have shown that genes such as *MET1* [80], *DRM1*, *DRM2* [55], *CMT3* [81, 82], *DDM1* [83], *DME* [84], *DEMETER-LIKE2* [85], and *ROS1* [86] play crucial roles in maintaining various genomic methylation states. To determine whether the methylation states correspond to different developmental stages in poplar stems, the orthologs of the above genes in *P. trichocarpa* were retrieved from the Phytozome database [87] using a Basic Local Alignment Search Tool (BLAST) program named BLASTP. *DNA methyltransferase (PtrMET1A/B)*, *DOMAINS REARRANGED METHYLTRANSFERASE 1 (PtrDRM1/2-A-C)*, *CHROMOMETHYLASE 3 (PtrCMT3-A-C)*, and *DECREASE IN DNA METHYLATION 1 (PtrDDM1-A/B)* are each involved in DNA methylation [88], and *DEMETER (DME) DNA demethylase (PtrDME-A/B)*, *PtrDEMETER-LIKE 2-A/B*, and *REPRESSOR OF SILENCING 1 (PtrROS1)* are involved in DNA demethylation. The expression levels of these genes in PS, TS, and SS were analyzed using qRT-PCR, as in our previous study [89]. Transcript of *PtrActin* was used as an endogenous control to normalize expression in different samples. Primers used for these studies are listed in Additional file 17. The expression level of each gene relative to the reference gene was calculated using the delta-delta CT method. Each measurement was carried out with three biological replicates, and each biological replicate has three technical replicates. Each error bar represents a standard error (SE) of the mean fold changes of the three biological replicates.

WGBS library construction and high-throughput sequencing

Plant samples were ground to powder in liquid nitrogen and genomic DNA was isolated using the DNeasy Plant Mini Kit (Qiagen China, Shanghai). The quantity and quality of extracted DNA were determined with a

NanoDrop ND-1000 (Thermo, USA). Then, genomic DNA was fragmented to 100-300 bp by sonication (Covaris, Massachusetts, USA) and purified with a Mini Elute PCR Purification Kit (Qiagen, MD, USA). The fragmented DNAs were end-repaired and a single “A” nucleotide was added to the 3’ end of the blunt fragments. Then the genomic fragments were ligated to methylated sequencing adapters. Adaptor-added DNA was subjected to bisulfite conversion using the EZ DNA Methylation Gold Kit (Zymo Research), and the bisulfite-treated DNA was PCR amplified for 16 cycles. The resultant DNA was subjected to paired-end sequencing on an Illumina HiSeq 2500 sequencer; reads of 150 nucleotides were generated.

Mapping and processing of BS-Seq reads

Raw read sequences generated by the Illumina pipeline in FastQ format were first subjected to quality control [90]. To get high quality clean reads, raw reads were filtered according to the following rules: 1) remove reads containing more than 10% unknown nucleotides (N); 2) remove low quality reads containing more than 40% low quality (Q-value \leq 20) bases. Clean reads were then mapped to the *P. trichocarpa* reference genome using BSMAP software (version: 2.90) with default parameter [91]. Briefly, both clean reads and the reference genome were first transformed into bisulfite-converted versions (C-to-T and G-to-A converted). Then, the converted reads were aligned to the similar versions of the converted genome in a directional manner. To eliminate the bias produced by the alignment of duplicates generated by PCR, a de-duplication step was used to remove reads mapping to the same position of the reference genome. Finally, sequence reads that produced a unique best alignment from the two alignment processes (original top and bottom strand) were compared to the normal genomic sequence, and the methylation state of all cytosine positions in the read were inferred [92]. The bisulfite conversion rate of lambda DNA was calculated and used as a measure of the false discovery rate in the identification of the methylation site according to the binomial probability distribution. Each replicate (sample) was mapped to genome separately. After read mapping, we merged multiple replicates by running a 200 bp sliding window in a step of 100 bp. The methylation rates of the same windows, which have the same chromosomal position, across different replicates were merged to get the unified methylated cytosine percentage. The methylation levels of each of the three methylation types, CG, CHG, and CHH, were calculated in the whole genome, in each chromosome, and in different genomic regions. For all genic regions, methylation profiling of gene bodies and 2 kb flanking regions (both sides of the genes) were plotted based on the average methylation levels of different sliding windows.

Identification of differentially methylated regions

Differential DNA methylation between the two samples at each locus was determined using Pearson's chi-square test (χ^2) in methyl Kit (version: 1.7.10) [93]. The minimum read coverage for a base to call a methylation status was set to 4. DMRs for each sequence context (CG, CHG, and CHH) conform to the following criteria: 1) For CG and CHG, numbers of GC or CHG in each window ≥ 5 , absolute value of the difference in methylation ratio ≥ 0.25 , and $q \leq 0.05$; 2) For CHH, numbers in a window ≥ 15 , absolute value of the difference in methylation ratio ≥ 0.15 , and $q \leq 0.05$; 3) For all C, numbers in a window ≥ 20 , absolute value of the difference in methylation ratio ≥ 0.2 , and $q \leq 0.05$.

Transcriptome sequencing and gene expression analysis

The same plant materials used for methylome analysis were also used for RNA isolation. The materials were first ground to powder in liquid nitrogen and total RNA was isolated using an RNA isolation kit (Auto Lab Biotechnology, Beijing, China). Using the RNase-Free DNase Set (Qiagen), we performed on-column DNase digestions three times during the RNA purification. The quantity and quality of extracted RNA were determined with a NanoDrop ND-1000 (Thermo, USA).

Whole transcriptome libraries were constructed using the NEB Next Ultra Directional RNA Library Prep Kit for Illumina (NEB, Ipswich, MA, USA) according to the manufacturer's instructions; the resulting libraries were assessed for size, quantitation, integrity and purity using a Bioanalyzer 2100 system and qPCR (Kapa Biosystems, Woburn, MA, USA). The libraries with good quality were subsequently sequenced on a HiSeq 2500 instrument that was set to produce 125 bp paired-end reads of 125 nucleotide long. Raw sequences were cleaned as follows: 1) remove reads containing adapters; 2) remove reads containing more than 10% unknown nucleotides (N); 3) remove low quality reads containing more than 50% low quality ($Q\text{-value} \leq 20$) bases. After that, the clean reads from all the samples were mapped to the *P. trichocarpa* genome using Bowtie2 [94] and TopHat2 [95] software with default parameters [96]. The expression levels of the protein-coding genes were calculated and normalized using fragments per kilobase of gene per million mapped fragments (FPKM) by Cufflinks (version 2.2.1) [97].

Genes were divided into non-expressed genes ($\text{FPKM} \leq 1$) and expressed genes. The latter were further divided into three groups based on expression levels, including minimally-expressed genes (Low, $1 < \text{FPKM} \leq 10$), moderately expressed genes (Moderate, $10 < \text{FPKM} \leq 100$), and highly expressed genes (High, $\text{FPKM} \geq 100$).

Relation between DNA methylation and gene expression

DMRs were annotated using the gene annotation file *P. trichocarpa_444_v3.1.gene.gff3.gz* from the *P. trichocarpa*

reference genome provided by Phytozome [78, 87]. The positions and types of TEs were obtained from *P. trichocarpa_533_v4.1.repeatmasked_assembly_v4.0.gff3.gz* downloaded from the Joint Genome Institute website (<https://genome.jgi.doe.gov/portal/pages/dynamic-OrganismDownload.jsf?organism=Phytozome>) [37]. Annotations were done by comparing the chromosome position information of DMRs with the corresponding annotation information in the gene annotation file. When a DMR overlapped a gene (> 1 nt), including a gene body and the 2 kb flanking region on either side, the DMR was associated with this gene and was then designated a DMG. Spearman correlation analysis was performed to discern statistical relationship between DNA methylation and gene expression within gene bodies and their ± 2 kb flanking regions. To explore the potential contribution of DNA methylation to the differentiation of gene expression, DMGs, DEGs, and common genes between DMGs and DEGs were subjected to KEGG pathway enrichment analysis (KEGG: <http://www.genome.jp/kegg/>). KOBAS software [98] was used to test the statistical enrichment of DEGs in various KEGG pathways.

Data analysis

The data were analyzed using one-way ANOVA and Duncan's multiple range test using SPSS 21 (Chicago, IL, USA). A statistically significant level was set to a p value < 0.05 . The data are presented as mean \pm standard error (SE) with each SE being calculated from three independent biological samples.

Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s12864-020-06902-6>.

Additional file 1. Summary of the methylation ratio in different sequence contexts (CG, CHG, and CHH) in developmental stems of *Populus trichocarpa*.

Additional file 2. Distributions of 5-methylcytosine density on Chromosomes 2–19 in primary stems (PS).

Additional file 3. Distributions of 5-methylcytosine density on Chromosomes 1–19 on transitional stems (TS).

Additional file 4. Distributions of 5-methylcytosine density on Chromosomes 1–19 on secondary stems (SS).

Additional file 5. Relative proportions of mCs in three sequence contexts (CG, CHG and CHH) in *Populus trichocarpa* stems.

Additional file 6. Comparison of mCs in CG, CHG, and CHH sequence contexts in three developmental stages: primary stems (PS), transitional stems (TS) and secondary stems (SS). The y-axis indicates the percentage of mC (%). The different letters on the top of the error bars indicate statistically significant differences between means at $p < 0.05$.

Additional file 7. Relative percentages of 5-mCs in three sequence contexts (CG, CHG, and CHH) in various transposable elements (TEs) of *Populus trichocarpa* stems.

Additional file 8. Methylation levels in various transposable elements (TEs) and their 2 kb proximal regions in primary stems (PS), transitional stems (TS), and secondary stems (SS). The y-axis represents methylation levels.

Additional file 9. Comparison of methylation levels in different developmental stages: primary stems (PS), transitional stems (TS), and secondary stems (SS). The y-axis represents methylation levels. The different letters on the top of the error bars indicate statistically significant differences between means at $p < 0.05$.

Additional file 10. Relationship between gene methylation (5-methylcytosines) and gene expression. (A), (B), and (C) Expression profiles of methylated genes compared with unmethylated genes in CG, CHG, and CHH contexts in primary stems (PS), respectively. (D), (E), and (F) Expression profiles of methylated genes compared with unmethylated genes in CG, CHG, and CHH contexts in transitional stems (TS), respectively. (G), (H), and (I) Expression profiles of methylated genes compared with unmethylated genes in CG, CHG, and CHH contexts in secondary stems (SS), respectively. The y-axis indicates the methylation levels. None represents the unmethylated group. Low represents the bottom third of the methylation group. Moderate represents the middle third of the methylation group. High represents the top third of the methylation group.

Additional file 11. Correlation between 5-methylcytosine methylation in different genic regions and gene expression in primary stems (PS), transitional stems (TS), and secondary stems (SS) of poplar. (A), (B), and (C) represent CG, CHG, and CHH DNA methylation contexts, respectively. The regions of promoter, gene body, and 2 kb downstream regions were split on the x-axis to investigate the Spearman rank correlation (y-axis) between levels of methylation and expression. $\rho > 0$ means positive correlation, and $\rho < 0$ means negative correlation.

Additional file 12. KEGG pathway enrichment of the rest of the differentially methylated genes (DMGs) and differentially expressed genes (DEGs). (A) and (B) represent KEGG pathway enrichment of the remaining DMGs and DEGs in primary stems (PS) vs transitional stems (TS), respectively. (C) and (D) represent KEGG pathway enrichment of the remaining DMGs and DEGs in TS vs secondary stems (SS), respectively. The size of the circle represents gene numbers, and the colors represents the q-value.

Additional file 13. The transcription factors (TFs) in primary stems (PS) vs transitional stems (TS).

Additional file 14. The transcription factors (TFs) in transitional stems (TS) vs secondary stems (SS).

Additional file 15. Shared genes in differentially expressed genes (DEGs) and differentially 5-methylcytosine methylated genes (DMGs) in primary stems (PS) vs transitional stems (TS).

Additional file 16. Shared genes in differentially expressed genes (DEGs) and differentially 5-methylcytosine methylated genes (DMGs) in transitional stems (TS) vs secondary stems (SS).

Additional file 17. Homologs of DNA methyltransferase and demethylase proteins in poplar.

Abbreviations

5mC: 5-methylcytosine; LTR: Long terminal repeats; IN: Internode; PS: Primary stems; TS: Transitional stems; SS: Secondary stems; WGBS: Whole-genome bisulfite sequencing; TEs: Transposable elements; LINE: Long interspersed nuclear elements; SINE: Short interspersed nuclear elements; DMRs: Differentially methylated regions; DMGs: Differentially methylated genes; DEGs: Differentially expressed genes; TFs: Transcription factors; SA: Salicylic acid; ABA: Abscisic acid; BLAST: Basic Local Alignment Search Tool

Acknowledgments

We thank Dr. Jennifer Sanders for proofreading this manuscript. Moreover, we also thank Guangzhou Genedenovo Biotechnology Co., Ltd. for assisting in sequencing and bioinformatics analysis.

Authors' contributions

YZ and CL: conducted most experiments and data analysis; HC and ST: participated in the vegetative propagation of *P. trichocarpa* plantlets; YL and SW: participated in histochemical staining; HZ and MS: wrote manuscript; HW: performed data analysis and wrote manuscript; ZW: designed the

experiments and wrote manuscript. All authors read and approved the final version of the manuscript.

Funding

The experimental design and implementation, RNA/DNA extraction, sequencing, data preprocessing analysis as well as interpretation were supported by the Fundamental Research Funds of Chinese Academy of Forestry (CAFYBB2018ZY001–5). The graduate students, Y. Z., C. L., H. C., and S. T., were financially supported by the Fundamental Research Funds for the Central Non-profit Research Institution of CAF (CAFYBB2018GB001). The graduate students, Y. L. and S. W., were supported by National Natural Science Foundation of China (31770640).

Availability of data and materials

The datasets used and analysed in this study are available in Sequence Read Archive (SRA) at NCBI at National Center for Biotechnology Information. The accession numbers for methylation data and RNA-seq data are PRJNA628812 and PRJNA628501, respectively. The data sets can be accessed at SRA (<https://www.ncbi.nlm.nih.gov/sra>) with the accession number provided above or <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA628812> and <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA628501>.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹State Key Laboratory of Tree Genetics and Breeding, Northeast Forestry University, Harbin, Heilongjiang 150040, People's Republic of China.

²Research Center of Saline and Alkali Land of State Forestry and Grassland Administration, Chinese Academy of Forestry, Beijing 100091, People's Republic of China. ³Institute of Soil and Environmental Sciences, University of Agriculture, Faisalabad 38000, Pakistan. ⁴College of Forest Resources and Environmental Science, Michigan Technological University, Houghton, MI 49931, USA.

Received: 13 March 2020 Accepted: 10 July 2020

Published online: 20 July 2020

References

- Plomion C, Leprovost G, Stokes A. Wood formation in trees. *Plant Physiol.* 2001;127(4):1513–23.
- Ye ZH, Zhong R. Molecular control of wood formation in trees. *J Exp Bot.* 2015;66(14):4119–31.
- Hertzberg M, Aspeborg H, Schrader J, Andersson A, Erlandsson R, Blomqvist K, Bhalerao R, Uhlen M, Teeri TT, Lundeberg J, et al. A transcriptional roadmap to wood formation. *Proc Natl Acad Sci U S A.* 2001;98(25):14732–7.
- Wang M, Qi X, Zhao S, Zhang S, Lu MZ. Dynamic changes in transcripts during regeneration of the secondary vascular system in *Populus tomentosa* Carr revealed by cDNA microarrays. *BMC Genomics.* 2009;10:215.
- Groover A, Robischon M. Developmental mechanisms regulating secondary growth in woody plants. *Curr Opin Plant Biol.* 2006;9(1):55–8.
- Baucher M, El Jaziri M, Vandeputte O. From primary to secondary growth: origin and development of the vascular system. *J Exp Bot.* 2007;58(13):3485–501.
- Demura T, Fukuda H. Transcriptional regulation in wood formation. *Trends Plant Sci.* 2007;12(2):64–70.
- Soler M, Plasencia A, Larbat R, Pouzet C, Jauneau A, Rivas S, Pesquet E, Lapiere C, Truchet I, Grima-Pettenati J. The Eucalyptus linker histone variant EgH1.3 cooperates with the transcription factor EgMYB1 to control lignin biosynthesis during wood formation. *New Phytol.* 2017;213(1):287–99.
- Wyatt GR. Recognition and estimation of 5-methylcytosine in nucleic acids. *Biochem J.* 1951;48(5):581–4.
- Doskocil J, Sorm F. Distribution of 5-methylcytosine in pyrimidine sequences of deoxyribonucleic acids. *Biochim Biophys Acta.* 1962;55:953–9.

11. Vaughn MW, Tanurdzic M, Lippman Z, Jiang H, Carrasquillo R, Rabinowicz PD, Dedhia N, McCombie WR, Agier N, Bulski A, et al. Epigenetic natural variation in *Arabidopsis thaliana*. *PLoS Biol.* 2007;5(7):e174.
12. Zilberman D, Gehring M, Tran RK, Ballinger T, Henikoff S. Genome-wide analysis of *Arabidopsis thaliana* DNA methylation uncovers an interdependence between methylation and transcription. *Nat Genet.* 2006;39(1):61–9.
13. Zhang X, Yazaki J, Sundaresan A, Cokus S, Chan WL, Chen H, Henderson IR, Shinn P, Pellegrini M, Jacobsen SE. Genome-wide high-resolution mapping and functional analysis of DNA methylation in *Arabidopsis*. *Cell.* 2006;126(6):1189–201.
14. Lafon-Placette C, Faivre-Rampant P, Delaunay A, Street N, Brignolas F, Maury S. Methylome of DNase I sensitive chromatin in *Populus trichocarpa* shoot apical meristematic cells: a simplified approach revealing characteristics of gene-body DNA methylation in open chromatin state. *New Phytol.* 2013;197(2):416–30.
15. Brautigam K, Soolanayakanahally R, Champigny M, Mansfield S, Douglas C, Campbell MM, Cronk Q. Sexual epigenetics: gender-specific methylation of a gene in the sex determining region of *Populus balsamifera*. *Sci Rep.* 2017;7:45388.
16. Song YP, Tian M, Ci D, Zhang DQ. Methylation of microRNA genes regulates gene expression in bisexual flower development in andromonoecious poplar. *J Exp Bot.* 2015;66(7):1891–905.
17. Schonberger B, Chen XC, Mager S, Ludewig U. Site-dependent differences in DNA methylation and their impact on plant establishment and phosphorus nutrition in *Populus trichocarpa*. *PLoS One.* 2016;11(12):e0168623.
18. Le Gac AL, Lafon-Placette C, Chauveau D, Segura V, Delaunay A, Fichot R, Marron N, Le Jan I, Berthelot A, Bodineau G, et al. Winter-dormant shoot apical meristem in poplar trees shows environmental epigenetic memory. *J Exp Bot.* 2018;69(20):4821–37.
19. Lafon-Placette C, Le Gac AL, Chauveau D, Segura V, Delaunay A, Lesage-Descauses MC, Hummel I, Cohen D, Jesson B, Le Thiec D, et al. Changes in the epigenome and transcriptome of the poplar shoot apical meristem in response to water availability affect preferentially hormone pathways. *J Exp Bot.* 2018;69(3):537–51.
20. Ding CJ, Liang LX, Diao S, Su XH, Zhang BY. Genome-wide analysis of day/night DNA methylation differences in *Populus nigra*. *PLoS One.* 2018;13(1):e0190299.
21. Gourcilleau D, Bogeat-Triboulot MB, Le Thiec D, Lafon-Placette C, Delaunay A, Abu El-Soud W, Brignolas F, Maury S. DNA methylation and histone acetylation: genotypic variations in hybrid poplars, impact of water deficit and relationships with productivity. *Ann Forest Sci.* 2010;67(2):208.
22. Law JA, Jacobsen SE. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat Rev Genet.* 2010;11(3):204–20.
23. Lister R, O'Malley RC, Tonti-Filippini J, Gregory BD, Berry CC, Millar AH, Ecker JR. Highly integrated Single-Base resolution maps of the Epigenome in *Arabidopsis*. *Cell.* 2008;133(3):523–36.
24. Cokus SJ, Feng S, Zhang X, Chen Z, Merriman B, Haudenschild CD, Pradhan S, Nelson SF, Pellegrini M, Jacobsen SE. Shotgun bisulphite sequencing of the *Arabidopsis* genome reveals DNA methylation patterning. *Nature.* 2008;452(7184):215–9.
25. Zemach A, Kim MY, Silva P, Rodrigues JA, Dotson B, Brooks MD, Zilberman D. Local DNA hypomethylation activates genes in rice endosperm. *Proc Natl Acad Sci U S A.* 2010;107(43):18729–34.
26. Kashkush K, Khasdan V. Large-scale survey of cytosine methylation of Retrotransposons and the impact of readout transcription from long terminal repeats on expression of adjacent Rice genes. *Genetics.* 2007;177(4):1975–85.
27. Soppe WJ, Jacobsen SE, Alonso-Blanco C, Jackson JP, Kakutani T, Koornneef M, Peeters AJ. The late flowering phenotype of *fwa* mutants is caused by gain-of-function epigenetic alleles of a homeodomain gene. *Mol Cell.* 2000;6(4):791–802.
28. Lippman Z, Gendrel AV, Black M, Vaughn MW, Dedhia N, McCombie WR, Lavine K, Mittal V, May B, Kasschau KD, et al. Role of transposable elements in heterochromatin and epigenetic control. *Nature.* 2004;430(6998):471–6.
29. Ma KF, Song YP, Yang XH, Zhang ZY, Zhang DQ. Variation in genomic methylation in natural populations of Chinese white poplar. *PLoS One.* 2013;8(5):e63977.
30. Vining KJ, Pomraning KR, Wilhelm LJ, Priest HD, Pellegrini M, Mockler TC, Freitag M, Strauss SH. Dynamic DNA cytosine methylation in the *Populus trichocarpa* genome: tissue-level variation and relationship to gene expression. *BMC Genomics.* 2012;13(1):27.
31. Wang Q, Ci D, Li T, Li P, Song Y, Chen J, Quan M, Zhou D, Zhang D. The role of DNA methylation in Xylogenesis in different tissues of poplar. *Front Plant Sci.* 2016;7:1003.
32. Prassinis C, Ko JH, Han KH. Transcriptome profiling of vertical stem segments provides insights into the genetic regulation of secondary growth in hybrid aspen trees. *Plant Cell Physiol.* 2005;46(8):1213–25.
33. van Raemdonck D, Pesquet E, Cloquet S, Beeckman H, Boerjan W, Goffner D, El Jaziri M, Baucher M. Molecular changes associated with the setting up of secondary growth in aspen. *J Exp Bot.* 2005;56(418):2211–27.
34. Stroud H, Greenberg MV, Feng S, Bernatavichute YV, Jacobsen SE. Comprehensive analysis of silencing mutants reveals complex regulation of the *Arabidopsis* methylome. *Cell.* 2013;152(1–2):352–64.
35. Li X, Zhu JD, Hu FY, Ge S, Ye MZ, Xiang H, Zhang GJ, Zheng XM, Zhang HY, Zhang SL, et al. Single-base resolution maps of cultivated and wild rice methylomes and regulatory roles of DNA methylation in plant gene expression. *BMC Genomics.* 2012;13:300.
36. Daccord N, Celton JM, Linsmith G, Becker C, Choise N, Schijlen E, van de Geest H, Bianco L, Micheletti D, Velasco R, et al. High-quality de novo assembly of the apple genome and methylome dynamics of early fruit development. *Nat Genet.* 2017;49(7):1099–106.
37. Pinosio S, Giacomello S, Faivre-Rampant P, Taylor G, Jorge V, Le Paslier MC, Zaina G, Bastien C, Cattonaro F, Marroni F, et al. Characterization of the poplar pan-genome by genome-wide identification of structural variation. *Mol Biol Evol.* 2016;33(10):2706–19.
38. Chodavarapu RK, Feng S, Ding B, Simon SA, Lopez D, Jia Y, Wang GL, Meyers BC, Jacobsen SE, Pellegrini M. Transcriptome and methylome interactions in rice hybrids. *Proc Natl Acad Sci U S A.* 2012;109(30):12040–5.
39. Zhang J, Liu Y, Xia EH, Yao QY, Liu XD, Gao LZ. Autotetraploid rice methylome analysis reveals methylation variation of transposable elements and their effects on gene expression. *Proc Natl Acad Sci U S A.* 2015;112(50):E7022–9.
40. Li Q, Eichten SR, Hermanson PJ, Zaunbrecher VM, Song J, Wendt J, Rosenbaum H, Madzima TF, Sloan AE, Huang J, et al. Genetic perturbation of the maize methylome. *Plant Cell.* 2014;26(12):4602–16.
41. Song QX, Lu X, Li QT, Chen H, Hu XY, Ma B, Zhang WK, Chen SY, Zhang JS. Genome-wide analysis of DNA methylation in soybean. *Mol Plant.* 2013;6(6):1961–74.
42. Zhong S, Fei Z, Chen YR, Zheng Y, Huang M, Vrebalov J, McQuinn R, Gapper N, Liu B, Xiang J, et al. Single-base resolution methylomes of tomato fruit development reveal epigenome modifications associated with ripening. *Nat Biotechnol.* 2013;31(2):154–9.
43. Ausin I, Feng S, Yu C, Liu W, Kuo HY, Jacobsen EL, Zhai J, Gallego-Bartolome J, Wang L, Egertsdotter U, et al. DNA methylome of the 20-gigabase Norway spruce genome. *Proc Natl Acad Sci U S A.* 2016;113(50):E8106–13.
44. Ong-Abdullah M, Ordway JM, Jiang N, Ooi SE, Kok SY, Sarpan N, Azimi N, Hashim AT, Ishak Z, Rosli SK, et al. Loss of karna transposon methylation underlies the mantled somaclonal variant of oil palm. *Nature.* 2015;525(7570):533–7.
45. Liang D, Zhang Z, Wu H, Huang C, Shuai P, Ye CY, Tang S, Wang Y, Yang L, Wang J, et al. Single-base-resolution methylomes of *Populus trichocarpa* reveal the association between DNA methylation and drought stress. *BMC Genet.* 2014;15(Suppl 1):S9.
46. Niederhuth CE, Bewick AJ, Ji LX, Alabady MS, Kim KD, Li Q, Rohr NA, Rambani A, Burke JM, Udall JA, et al. Widespread natural variation of DNA methylation within angiosperms. *Genome Biol.* 2016;17:194.
47. Xu JD, Zhou SS, Gong XQ, Song Y, van Nocker S, Ma FW, Guan QM. Single-base methylome analysis reveals dynamic epigenomic differences associated with water deficit in apple. *Plant Biotechnol J.* 2018;16(2):672–87.
48. Liang D, Zhang Z, Wu H, Huang C, Shuai P, Ye CY, Tang S, Wang Y, Yang L, Wang J. Single-base-resolution methylomes of *Populus trichocarpa* reveal the association between DNA methylation and drought stress. *BMC Genet.* 2014;15(1 Supplement):S9.
49. Seymour DK, Koenig D, Hagmann J, Becker C, Weigel D. Evolution of DNA methylation patterns in the Brassicaceae is driven by differences in genome organization. *PLoS Genet.* 2014;10(11):e1004785.
50. Li SF, Guo YJ, Li JR, Zhang DX, Wang BX, Li N, Deng CL, Gao WJ. The landscape of transposable elements and satellite DNAs in the genome of a dioecious plant spinach (*Spinacia oleracea* L.). *Mob DNA.* 2019;10:3.
51. Du Q, Wang B, Wei Z, Zhang D, Li B. Genetic diversity and population structure of Chinese white poplar (*Populus tomentosa*) revealed by SSR markers. *J Hered.* 2012;103(6):853–62.
52. Ma KF, Song YP, Jiang XB, Zhang ZY, Li BL, Zhang DQ. Photosynthetic response to genome methylation affects the growth of Chinese white poplar. *Tree Genetics Genome.* 2012;8:1407–21.

53. Su YT, Bai XT, Yang WL, Wang WW, Chen ZY, Ma JC, Ma T. Single-base-resolution methylomes of *Populus euphratica* reveal the association between DNA methylation and salt stress. *Tree Genet Genomes*. 2018;14(6):86.
54. Bartels A, Han Q, Nair P, Stacey L, Gaynier H, Mosley M, Huang QQ, Pearson JK, Hsieh TF, An YC, et al. Dynamic DNA methylation in plant growth and development. *Int J Mol Sci*. 2018;19(7):2144.
55. Matzke MA, Mosher RA. RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. *Nat Rev Genet*. 2014;15(6):394–408.
56. Chan SW, Henderson IR, Jacobsen SE. Gardening the genome: DNA methylation in *Arabidopsis thaliana*. *Nat Rev Genet*. 2005;6(5):351–60.
57. Finnegan EJ, Peacock WJ, Dennis ES. DNA methylation, a key regulator of plant development and other processes. *Curr Opin Genet Dev*. 2000;10(2):217–23.
58. Kilby NJ, Leyser HM, Furner IJ. Promoter methylation and progressive transgene inactivation in *Arabidopsis*. *Plant Mol Biol*. 1992;20(1):103–12.
59. Weinhold A, Kallenbach M, Baldwin IT. Progressive 35S promoter methylation increases rapidly during vegetative development in transgenic *Nicotiana attenuata* plants. *BMC Plant Biol*. 2013;13:99.
60. Suzuki Y, Arae T, Green PJ, Yamaguchi J, Chiba Y. AtCCR4a and AtCCR4b are involved in determining the poly(a) length of granule-bound starch synthase 1 transcript and modulating sucrose and starch metabolism in *Arabidopsis thaliana*. *Plant Cell Physiol*. 2015;56(5):863–74.
61. Korolev AV, Chan J, Naldrett MJ, Doonan JH, Lloyd CW. Identification of a novel family of 70 kDa microtubule-associated proteins in *Arabidopsis* cells. *Plant J*. 2005;42(4):547–55.
62. Woo HH, Jeong BR, Hirsch AM, Hawes MC. Characterization of *Arabidopsis* AtUGT85A and AtGUS gene families and their expression in rapidly dividing tissues. *Genomics*. 2007;90(1):143–53.
63. Wu Y, Hou J, Yu F, Nguyen STT, McCurdy DW. Transcript profiling identifies NAC-domain genes involved in Regulating Wall ingrowth deposition in phloem parenchyma transfer cells of *Arabidopsis thaliana*. *Front Plant Sci*. 2018;9:341.
64. Kao YY, Harding SA, Tsai CJ. Differential expression of two distinct phenylalanine ammonia-lyase genes in condensed tannin-accumulating and lignifying cells of quaking aspen. *Plant Physiol*. 2002;130(2):796–807.
65. Brown DM, Zeef LA, Ellis J, Goodacre R, Turner SR. Identification of novel genes in *Arabidopsis* involved in secondary cell wall formation using expression profiling and reverse genetics. *Plant Cell*. 2005;17(8):2281–95.
66. Shi R, Ying-Hsuan S, Li Q, Heber S, Sederoff R, Chiang VL. Towards a systems approach for lignin biosynthesis in *Populus trichocarpa*: transcript abundance and specificity of the Monolignol biosynthetic genes. *Plant Cell Physiol*. 2010;51(1):144–63.
67. Zhong R, Lee C, Zhou J, McCarthy RL, Ye ZH. A battery of transcription factors involved in the regulation of secondary Cell Wall biosynthesis in *Arabidopsis*. *Plant Cell*. 2008;20(10):2763–82.
68. Eulgem T, Somssich IE. Networks of WRKY transcription factors in defense signaling. *Curr Biol*. 2007;10(4):366–71.
69. Jackson RG, Kowalczyk M, Li Y, Higgins G, Ross J, Sandberg G, Bowles DJ. Over-expression of an *Arabidopsis* gene encoding a glucosyltransferase of indole-3-acetic acid: phenotypic characterisation of transgenic lines. *Plant J*. 2002;32(4):573–83.
70. Mito T, Seki M, Shinozaki K, Ohme-Takagi M, Matsui K. Generation of chimeric repressors that confer salt tolerance in *Arabidopsis* and rice. *Plant Biotechnol J*. 2011;9(7):736–46.
71. Tian S, Wang X, Li P, Wang H, Ji H, Xie J, Qiu Q, Shen D, Dong H. Plant aquaporin AtPIP1.4 links Apoplastic H₂O₂ induction to disease immunity pathways. *Plant Physiol*. 2016;171(3):1635–50.
72. Yasumoto S, Fukushima EO, Seki H, Muranaka T. Novel triterpene oxidizing activity of *Arabidopsis thaliana* CYP716A subfamily enzymes. *FEBS Lett*. 2016;590(4):533–40.
73. Tsai CJ, Harding SA, Tschaplinski TJ, Lindroth RL, Yuan Y. Genome-wide analysis of the structural genes regulating defense phenylpropanoid metabolism in *Populus*. *New Phytol*. 2006;172(1):47–62.
74. Sterky F, Bhalerao RR, Unneberg P, Segerman B, Nilsson P, Brunner AM, Campaa LC, Lindvall JJ, Tandré K, Strauss SH. A *Populus* EST resource for plant functional genomics. *Proc Natl Acad Sci U S A*. 2004;101(38):13951–6.
75. Hamberger B, Ellis M, Friedmann M, Clarice DAS, Barbazuk B, Douglas CJ. Genome-wide analyses of phenylpropanoid-related genes in *Populus trichocarpa*, *Arabidopsis thaliana*, and *Oryza sativa*: the *Populus* lignin toolbox and conservation and diversification of angiosperm gene families. *Can J Bot*. 2007;85(12):1182–201.
76. Lu S, Zhou Y, Li L, Chiang VL. Distinct roles of cinnamate 4-hydroxylase genes in *Populus*. *Plant Cell Physiol*. 2006;47(7):905–14.
77. Downen RH, Pelizzola M, Schmitz RJ, Lister R, Downen JM, Nery JR, Dixon JE, Ecker JR. Widespread dynamic DNA methylation in response to biotic stress. *Proc Natl Acad Sci U S A*. 2012;109(32):E2183–91.
78. Tuskan GA, Difazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, Putnam N, Ralph S, Rombauts S, Salamov A, et al. The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science*. 2006;313(5793):1596–604.
79. Li S, Zhen C, Xu W, Wang C, Cheng Y. Simple, rapid and efficient transformation of genotype Nisqualy-1: a basic tool for the first sequenced model tree. *Sci Rep*. 2017;7(1):2638.
80. Lindroth AM, Cao X, Jackson JP, Zilberman D, McCallum CM, Henikoff S, Jacobsen SE. Requirement of CHROMOMETHYLASE3 for maintenance of CpXpG methylation. *Science*. 2001;292(5524):2077–80.
81. Lindroth AM, Shultis D, Jasencakova Z, Fuchs J, Johnson L, Schubert D, Patnaik D, Pradhan S, Goodrich J, Schubert I, et al. Dual histone H3 methylation marks at lysines 9 and 27 required for interaction with CHROMOMETHYLASE3. *EMBO J*. 2004;23(21):4286–96.
82. Jackson JP, Lindroth AM, Cao X, Jacobsen SE. Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. *Nature*. 2002;416(6880):556–60.
83. Jeddeloh JA, Stokes TL, Richards EJ. Maintenance of genomic methylation requires a SWI2/SNF2-like protein. *Nat Genet*. 1999;22(1):94–7.
84. Choi Y, Gehring M, Johnson L, Hannon M, Harada JJ, Goldberg RB, Jacobsen SE, Fischer RL. DEMETER, a DNA glycosylase domain protein, is required for endosperm gene imprinting and seed viability in *Arabidopsis*. *Cell*. 2002;110(1):33–42.
85. Penterman J, Zilberman D, Huh JH, Ballinger T, Henikoff S, Fischer RL. DNA demethylation in the *Arabidopsis* genome. *Proc Natl Acad Sci U S A*. 2007;104(16):6752–7.
86. Gong Z, Morales-Ruiz T, Ariza RR, Roldan-Arjona T, David L, Zhu JK. ROS1, a repressor of transcriptional gene silencing in *Arabidopsis*, encodes a DNA glycosylase/lyase. *Cell*. 2002;111(6):803–14.
87. Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, Fazo J, Mitros T, Dirks W, Hellsten U, Putnam N, et al. Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res*. 2012;40(Database issue):D1178–86.
88. Vining K, Pomraning KR, Wilhelm LJ, Ma C, Pellegrini M, Di Y, Mockler TC, Freitag M, Strauss SH. Methylome reorganization during in vitro dedifferentiation and regeneration of *Populus trichocarpa*. *BMC Plant Biol*. 2013;13:92.
89. Li M, Wang S, Liu Y, Zhang Y, Ren M, Liu L, Lu T, Wei H, Wei Z. Overexpression of PsnSuSy1, 2 genes enhances secondary cell wall thickening, vegetative growth, and mechanical strength in transgenic tobacco. *Plant Mol Biol*. 2019;100(3):215–30.
90. Chen S, Zhou Y, Chen Y, Gu J. Fastq: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*. 2018;34(17):i884–90.
91. Xi Y, Li W. BSMAP: whole genome bisulfite sequence MAPPING program. *BMC Bioinformatics*. 2009;10(1):232–0.
92. Lister R, Pelizzola M, Downen RH, Hawkins RD, Hon G, Tonti-Filippini J, Nery JR, Lee L, Ye Z, Ngo QM. Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature*. 2009;462(7271):315–22.
93. Akalin A, Kormaksson M, Li S, Garrett-Bakelman FE, Figueroa ME, Melnick A, Mason CE. methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome Biol*. 2012;13(10):R87.
94. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol*. 2009;10(3):R25.
95. Trapnell C, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics*. 2009;25(9):1105–11.
96. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol*. 2013;14(4):R36.
97. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ, Pachter L. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol*. 2010;28(5):511–5.
98. Xie C, Mao X, Huang J, Ding Y, Wu J, Dong S, Kong L, Gao G, Li CY, Wei L. KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. *Nucleic Acids Res*. 2011;39:W316–22.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.