Integrated analysis of full-length transcriptome and secondary metabolism reveals novel regulation in *Paulownia fortunei* infected with phytoplasma

Xiaogai Zhao

Henan Agricultural University

Xiaoqiao Zhai

Forestry Academy Henan

Zhe Wang

Henan Agricultural University

Yabing Cao

Henan Agricultural University

Zhenli Zhao

Henan Agricultural University

Xibing Cao

Henan Agricultural University

Guoqiang Fan (✉ zlx64@126.com)

Institute of Paulownia <https://orcid.org/0000-0003-0385-2578>

Research article

Keywords: *P. fortunei*; Full-length transcriptome; PaWB; Pfwuschel; Secondary metabolism; Lignin; Flavonoids

Posted Date: May 23rd, 2019

DOI: <https://doi.org/10.21203/rs.2.9770/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Background: Paulownia witches' broom (PaWB) is a devastating disease caused by phytoplasma, which can lead to considerable economic losses. Previous studies have shown that 60 mg•L⁻¹ methyl methane sulfonate (MMS) can restore infected Paulownia seedlings to healthy one. To improve the understanding of PaWB, we used single molecule real time (SMRT) sequencing, which can provide long sequences that can be effectively used to identify gene alternative splicing, obtain novel genes, improve gene annotations and comprehensive analysis of the metabolome. **Results:** Here, the first integrated analysis SMRT sequencing, the Illumina platforms, and high-performance liquid chromatography-mass spectrometry (HPLC-MS) to analyze changes in transcripts and metabolites in Paulownia fortunei. A total of 140,528 non redundant transcripts were identified, which supported 6,209 of novel gene loci that can enrich 70,223 annotation information of Paulownia fortunei genome. A total of 5,606 transcripts, 46 lncRNAs, 83 fusion transcripts, 597 transcription factors, and 57 metabolites were related indirectly or directly to PaWB. Subsequent analysis of the transcriptome and metabolome found trans-cinnamate 4-monooxygenase, caffeoyl-CoA-O methyltransferase, ferulic acid and peroxidase were up-regulated in Paulownia fortunei infected with phytoplasma. They were involved in flavonoid, phenylpropane and salicylic acid metabolism, which might be related to Paulownia-phytoplasma interaction. **Conclusions:** Our results will enrich the understanding of molecular mechanisms of PaWB. This study provides a foundation for better understanding the changes in gene expression, metabolites, and morphology of Paulownia fortunei with PaWB.

Background

Paulownia trees have been planted worldwide and play important roles in safeguarding food security (relay intercropping), the production of craft products (musical instruments and paper making), and in improving the ecological environment [1, 2]. However, Paulownia is easily infected with phytoplasma that causes Paulownia witches' broom (PaWB), leading to slow growing of big trees and death of small trees, which can lead to considerable economic losses [3].

PaWB has been studied for decades, with focus on morphological, physiological, and molecular changes [4]. It has been revealed that, in phytoplasma-infected seedlings, the leaves atrophy shrinking, axillary bud proliferation, shortened internodes, and phyllody, compared with their content in uninfected seedlings [5, 6]. From the physiological aspects, we obtained auxin, and amino acids decrease [7, 8], whereas cystine catalase, auxin oxidase, and vitamin C oxidase increase [8, 9]. In addition, we identified genes, proteins, and metabolites that responded to phytoplasma infection [10-16]. However, short-read sequence needs to be assembled into full-length transcripts, which lose a lot of alternative splicing (AS). Despite all these studies, a full understanding of PaWB is still lacking. To improve the understanding of PaWB, we used SMRT sequencing, which can provide long sequences of up to 20 kb that can be used effectively to identify gene alternative splicing, obtain novel genes, and improve gene annotations.

Studying metabolites can help to reveal changes in their complex interactions in response to biotic and abiotic stresses and establish direct contact with biological events [17, 18]. Comprehensive data from SMRT sequencing have been applied to study the functions of transcripts in regulating metabolism and environmental responses in plants, including *Camellia sinensis* (tea) [19], and *Archilochus colubris* (hummingbird) [20]. Hence,

a comprehensive analysis of the metabolome and transcriptome is a useful approach for identifying transcripts involved in PaWB.

To further elucidate the mechanism of PaWB in *P. fortunei*, we performed SMRT sequencing, sequencing on an Illumina platform, and high-performance liquid chromatography-mass spectrometry (HPLC-MS) to analyze changes in the transcripts and metabolites of healthy seedlings (PF), phytoplasma-infected seedlings (PFI), healthy seedlings treated with 60 mg•L⁻¹ methyl methane sulfonate (MMS) (PF-60), and phytoplasma-infected seedlings treated with 60 mg mg•L⁻¹ MMS (PFI-60). The first integrated analysis of the full-length transcriptome and **secondary metabolism** revealed transcripts involved in PaWB. This study will help in better understanding the changes in genes, metabolites, and morphology in phytoplasma-infected *P. fortunei*, and the results may be applied to other plants.

Results

SMRT sequencing and quality testing

We sequenced the transcriptomes of the PF, PFI, PF-60, and PFI-60 samples using SMRT sequencing platform. To avert loading bias of SMRT sequencing, cDNA insert size with 1-2 kb, 2-3 kb, and 3-6 kb were prepared, respectively. A total of 46.79 Gb clean data were obtained, of which, read of insert (ROI) contents of 276,222 (PF), 332,403 (PFI), 311,487 (PF-60), and 384,357 (PFI-60), and full-length non-chimeric reads (containing a poly (A) tail, and the 5' and 3' barcoded primers) were identified as follows: 124,712 (PF), 167,754 (PFI), 140,707 (PF-60), and 189,122 (PFI-60) (Table 1). After cluster analysis, **error** correction, and redundancy elimination, we obtained 140,528 transcript isoforms, including 28,709 (PF), 38,309 (PFI), 35,713 (PF-60), and 37,797 (PFI-60). The integrity of the sequences (Figure. 1a), the distribution of the dispersion degrees of the transcriptome expression levels (Figure. 1b) and the correlation analysis of the samples (Figure. 1c) were shown in Figure. 1, which demonstrated the SMRT sequencing quality was good enough to be used in the subsequent analysis.

To further validate the accuracy of the full-length transcriptome, six transcripts (selected from Table S1) of PaWB were selected randomly for quantitative real-time polymerase chain reaction (qRT-PCR) analysis. The expression patterns of qRT-PCR showed similar trends to the expression patterns obtained by analyzing the SMRT sequencing data (Figure. 1d). QRT-PCR results can support the correctness of SMRT sequencing.

Alternative splicing analysis

By aligning the nonredundant isoforms to the reference Paulownia genome (<http://paulownia.genomics.cn>), we found more exons from PacBio Iso-Seq than from the reference database (Figure. 2b). A total of 28,709 (PF), 38,309 (PFI), 35,713 (PF-60) and 37,797 (PFI-60) isoforms were identified, which were classified into three groups (A, B, and C, respectively) by performing alternative splices. Group A contained 13,082 known isoforms (detected and mapped to the gene set); group B contained 70,223 new isoforms (newly annotated genes); and group C contained 108,962 new isoforms from 3,386 likely novels genetic, which do not find any annotated gene (Figure. 2c). These novel gene loci improved 70,223 novel transcript annotations, among which five main AS types were identified. The numbers of AS events in PFI, PF-60, and PFI-60 were **more** than in PF. The most frequent AS event in the four samples was intron retention (IR) (Figure. 2a), which is in line with the results in other plant species [21]. IR can vary the open reading frames, resulting in functionally different variants.

Three generations of full-length transcriptome data compared with two generations of the transcriptome, more alternatively spliced transcripts were found, such as growth-regulating factor 2-like (Paulownia_LG9G000183.1, 3 isoforms) and 14-3-3 protein 1-like (Paulownia_LG5G000267.1, 2 isoforms). To further validate the accuracy of the detected isoforms, we randomly selected six genes, and verified the isoforms by real-time polymerase chain reaction (RT-PCR) (Table S2, Figure. 2d). The results showed that six transcripts each had two isoforms.

Long non-coding RNAs identification

Long non-coding RNAs (lncRNAs) response to biotic and abiotic stress, which changes in the expression levels in numerous organisms [22]. In this study, a total of 1,658 lncRNA transcripts were detected in the four samples, 301 (PF), 384 (PFI), 348 (PF-60), and 356 (PFI-60). According to the methods was referred to Wang [23], we identified 46 lncRNAs (18 lncRNAs down-regulation and 28 lncRNAs up-regulation, Table S3) that likely were associated with PaWB. Among them, 1,541 lncRNAs were classified them into four (A, B, C and D, respectively) groups according to their positions in <http://paulownia.genomics.cn> annotations. A contents of 132 (8.57%) antisense strand, B contents of 114 (7.40%) intronic regions, C contents of 529 (34.33%) intergenic regions and D contents of 766 (49.71%) sense strand (Figure. 3a, Figure. 4). Furthermore, 1,886 target genes prediction was performed. From that, the gene ontology (GO)functional classification and functional annotation of 1,886 differentially expressed target genes indicated that these target genes were involved in 9 GO classification (Figure. 3b), mainly including plant hormone signal transduction, secondary metabolism, and RNA transport.

Fusion transcript analysis

Fusion transcript is chimeric RNA encoded, which stem from transcription read-through, genomic structural rearrangements, or trans-splicing RNA [24, 25]. Gene fusion has a common feature by somatic chromosomal rearrangement in humans. However, gene fusion has a few reports in plants. In this study, we identified 93 fusion transcripts (81 in PF, 90 in PFI, 80 in PF-60 and 93 in PFI-60 (Table S4, Figure 4). The Venn diagram showed the highest number of fusion transcripts was found in all four samples, and 83 fusion transcripts were considered as tissues specific fusion transcripts in PFI (Figure S1).

Genes analysis

To recognize expression changes in response to phytoplasma infection in *P. fortunei*, initially, the fragments per kilobase of transcript per Million fragments mapped (FPKMs) of the selected sample genes were standardized and normalized. Then, the standardized FPKM values were analyzed by k-means clustering. We identified 30,437 genes, which were divided into 10 clusters (K1–K10, Figure. 5a). In K2 and K6, the genes were candidate PaWB-related genes (Table S5). To obtain more biological information about the differentially expressed genes, we conducted a GO enrichment analysis (Figure. 4b). In K2, we found that genes involved in the membrane were enriched in PFI. In K6, we found that genes involved in the membrane and cellular processes were enriched in PFI. Such as, homeobox gene WUSCHEL (WUS) can regulate axillary bud formation or branching [26, 27]. In this study, we found that Paulownia_LG5G000525 encoding protein WUS-like (PfWUS) were up-expressed in the PFI. This report is in line with the results in soybean [26], which may regulate [witches' broom](#) to promote axillary meristem and initiation the cytokinin (CK) signaling pathway [28].

Transcription factor analysis

Transcription factors (TFs) are proteins that bind specific nucleotide sequences on the upstream of the gene, which regulates RNA polymerase binding to DNA templates, thus regulating gene transcription. Previous studies demonstrate that many TFs related directly or indirectly to PaWB have been identified (Li et al., 2018). Hence, we analyzed the expression patterns of TFs in PF, PFI, PF-60, and PFI-60. A total of 12,162 TFs transcripts were obtained in different gradient (Figure 6a). Subsequent, weighted correlation network analysis (WGCNA) was used to identify correlation networks of the TFs. Among them, yoyalblue module of highly expressed were identified in PFI (Figure S2). We identified 597 (Table S6) PaWB-associated transcription factors. Based on the GO enrichment analysis, we found that transcripts involved in cellular process, catalytic activity, binding, and metabolic process were enriched may be important in PaWB (Figure S3).

Metabolites analysis

A total of 645 metabolites were detected, which mainly include plant hormones, sugars, and flavonoids (Table S7, Figure. 7a). To ascertain the repeatability of metabolites, different samples were analyzed for overlap analysis by Quality Control (QC) mass spectrometry. The results showed that the metabolites detected high total ion current curve overlap, PF, PFI, PF-60, and PFI-60 samples showed significant separation trend, indicating significant differences in metabolites between the four groups of samples (Figure. 7b). $R^2X=0.762$ and $Q^2=0.567$ main components were obtained in this study. The two QC almost completely coincided, indicating that the paulownia sample mass spectrometry detection analysis is more stable. Data repeatability and credibility is high, can be used for subsequent analysis (Figure. 7c). To recognize metabolite changes in response to the phytoplasma infection, initially, the fragments per kilobase of metabolite per Million fragments mapped (FPKMs) of the selected sample genes were standardized and normalized. For further analysis, we classified the metabolites of five comparisons (A, B, C, D, and E, respectively). In the A (PF-60 vs PF), B (PFI-60 vs PFI) and C (PFI vs PF) comparisons, we found 102 (44 up, 58 down), 94 (70 up, 24 down) and 109 (51 up, 58 down) differentially expressed secondary metabolites were identified, respectively. According to this reasonable comparison scheme (Fig 2), We detected 144 metabolites from comparison D (different in A and B), 57 metabolites from comparison E (same in D and C). Finally, 57 metabolites were selected as candidate PaWB-related metabolites, these metabolites were mainly related to flavonoid, phytohormone, and phenols (Table S8). When plants infected with phytoplasma, plant active oxygen to activate the plant's defense system, while too much reactive oxygen species can cause damage to the plant cells and the structure of genes. Flavonoids, as a kind of antioxidants, play an important role in this process. At the same time, the contents of IAA and its chelates were significantly different in phytoplasma-infected seedlings. This is consistent with our previous research results [29], speculating that the PaWB may be related to the change of auxin content. These results showed that phytoplasma infection activates the plant's defensive response and disrupts plant metabolism, further inducing an increase in antioxidant content and imbalance in plant hormones.

Cross-talk analysis between full-length transcript and metabolome

To explore the impact of phytoplasma infection in *P. fortunei*, we conducted a correlation analysis of the transcriptome and metabolome. In particular, ferulic acid (PT0417) correlated with peroxidase (PB.5202) in the phenylpropanoid biosynthesis pathway (Figure. 8d). Peroxidase is the last key enzyme in lignin synthesis, which plays an important role in plant biotic and abiotic stresses, and can regulate cell lignification. Similarly, ferulic acid (PT0417) correlated with peroxidase (PB.5202) in the phenylpropanoid biosynthesis pathway. In addition, in the phenylpropanoid biosynthesis pathway, we found that the gene (Paulownia_LG12G000996.1) encoding

trans-cinnamate 4-monooxygenase, a key enzyme in flavonoids biosynthesis, was up-regulated in PFI and that changes in its expression levels directly affected downstream compound synthesis, which is consistent with previous results [29]. Full-length transcriptome and secondary metabolism analysis revealed a novel regulation mechanism that will help enrich the understanding of Paulownia-phytoplasma interactions.

Discussion

Next-generation sequencing approaches are highly useful in quantifying gene expression [30] and have been used widely in transcriptome studies. However, short reads can lead to misassembly [31, 32], which prevents the construction of accurate full-length transcripts. SMRT sequencing, which can generate reads up to 20 kb in length, can effectively identify AS, promote complex transcription analysis, and identify new genes. SMRT sequencing has been used for a number of plants, including sorghum, maize, cotton, and populus [21, 30, 33, 34]. Therefore, combining next-generation sequencing and SMRT sequencing can improve the quality of sequencing, the identification and annotation of new genes, and provide new insights into gene expression and AS isomers. HPLC-MS can generate more metabolite profiles and has been used to analyze thermolabile and low concentration compounds [35]. In this study, we combined SMRT sequencing, sequencing on an Illumina platform, and HPLC-MS to study transcript of phytoplasma infection in *P. fortunei*. Full-length transcriptome and metabolomics analyses of the differentially expressed genes revealed that AS, LncRNAs, TF, fusion transcripts, metabolites, and secondary metabolites were associated indirectly or directly with phytoplasma infection in *Paulownia fortunei*. Closely related transcripts of witches' broom are responding to phytoplasma infection mainly from three aspects (i) the formation of physical barriers, such as the increase of lignin synthesis; (ii) abduction the production of disease-resistant transcription; (iii) cause the imbalance of endogenous plant hormone.

SMRT sequencing improves paulownia annotation

To improve the understanding of increases the accuracy of transcript, we used SMRT sequencing, which can provide long sequences that can be used effectively to identify gene alternative splicing, obtain novel genes, and improve gene annotations. We sequenced the transcriptomes of twelve mRNA samples from PF, PFI, PF-60, and PFI-60 samples. A total of 1,274,765,160 clean reads were mapped to paulownia reference genome, which 1,121,793,341 reads (87.81%) mapped in paulownia reference genome (Table S4). We quantified 286,907 consensus isoforms from twelve samples. The data of biological repeats were clustered together by clustering analysis (Table S5). A total of 140,528 nonredundant transcripts were identified, which supported 6,209 of novel gene loci that can enrich 70,223 annotation information of *Paulownia fortunei* genome (Figure. 2c).

Cross-talk between ferulic acid and peroxidase associated with PaWB

Plant responses to stress signals do not occur in isolation, rather different signaling pathways cross-talk to combat and tolerate the stress, and a signal can lead to alterations in genes and gene products. A transcriptome is considered to represent all genes in a cell at a certain time, whereas a metabolome represents the endpoint of gene expression, so differences in metabolites can be linked directly to biological events [17]. Integrative transcriptomic and metabolomic studies can help to identify key genes and metabolites associated with biotic and abiotic stresses and have been proved to be important in revealing metabolic regulation and identifying key candidate genes. In this study, several key genes and metabolites were identified, and the correlation between

ferulic acid (PT0417) and peroxidase (PB.5202) was revealed in the phenylpropanoid biosynthesis pathway. Peroxidase catalyzes the last step in lignin biosynthesis, and its activity and isozymes are related closely to disease resistance of plants [36]. When plants are infected by pathogenic bacteria, they will make lots of thick cell walls such as lignin [37], which constitute an important structural barrier. Similarly, ferulic acid cross-linked mainly with cell wall lignin and polysaccharides to form part of the cell wall, which solidifies the plant wall in the defense against pathogen intrusion [38]. In addition, ferulic acid is a good antioxidant that has strong scavenging effects on hydrogen peroxide, superoxide radicals, hydroxyl radicals, and nitroso peroxide, to protect normal metabolism [39]. Ferulic acid, caffeoyl-CoA-O-methyltransferase and peroxidase were up-regulated in PFI. These were beneficial to lignin production, which phenylpropane metabolism was enhanced to resist phytoplasma-infection Paulownia.

Flavonoids responsive to phytoplasma infection in *P. fortunei*

Flavonoids play important roles in pathogen-infected plants because they have antioxidant properties that prevent the breakdown of lipids by peroxidase [37]. Flavonoid compounds through the phenylpropanoid metabolic pathway, which uses phenylalanine and tyrosine as raw materials [40, 41]. These amino acids enter downstream branching pathways and the flavonoids are synthesized by phenylalanine ammonia-lyase, trans-cinnamate 4-monooxygenase, and cinnamate-4-hydroxylase catalyzes. In our full-length transcriptome study, we found that the gene (Paulownia_LG12G000996.1) encoding trans-cinnamate 4-monooxygenase, a key enzyme in flavonoids biosynthesis, was up-regulated in PFI and its changes in its expression levels directly affected the downstream compound synthesis, which is consistent with previous result [29]. These was beneficial to flavonoid production, which phenylpropane metabolism was enhanced to resist phytoplasma-infection Paulownia.

Paulownia may produce a large amount of active oxygen and activate the defense system after phytoplasma infection [9]. Flavonoids and other substances are used mainly as antioxidants to eliminate reactive oxygen species produced by plants in response to biological stress [42]. In *P. fortunei*, the flavonoids may be used as antioxidants to eliminate the damage caused by excess reactive oxygen, which may play a role in balancing plant defense responses.

Pfwus responded to PaWB to regulate morphological variations

When invaded by a pathogen, host plants initiate pathogen-associated molecular pattern-triggered immunity and effector-triggered immunity responses, and gene expression patterns of the host plant change [43]. Previous studies showed that squamosa promoter binding protein-like 9 (SPL9) is the major target of miR156b, and SPL9 by physically interacting with the WUS may regulate axillary bud formation and branching [26]. Notably, the WUS can directly regulate axillary bud formation and branching [23, 26], and the cytokinin (CK) signaling pathway [28]. WUSs have been reported in many model plants. Arabidopsis response regulator genes, which LATERAL ORGAN BOUNDARIES1 (WUS initiation) and BRANCHED1 (lateral bud outgrowth) have been identified [44-46]. Rice response regulator genes, which WUS orthologs (TILLERS ABSENT1 and MONOCUM 3) regulate tillering [46, 47]. In our study, we found the gene Paulownia_LG5G000525.1 (*Pfwus*) were up-regulated in the PFI. Notably, regarding *Pfwus* bases, the similarity with Arabidopsis (74%) and soybean (72%). *Pfwus* may directly regulate [witches' broom](#) to promote axillary meristem initiation, but further studies are needed.

WUS is a positive regulator of stem cells, which represses the transcription of two-component ARABIDOPSIS RESPONSE REGULATOR (ARR) genes (such as ARR6, ARR7, and ARR15) [48-50]. Takatoshi et al. indicated that CK signal transduction was negative regulated by ARR in Arabidopsis thaliana, which CK signal transduction was regulated by ARR [48](Figure. S4). In our study, we found *Pfwus* genes were up-regulated and two-component response regulator ARR-like (Paulownia_TIG00016041G000111.1) were down-regulated in the PFI. On the other hand, we results shown that zein O-glycosyltransferase-like (PB.21605.1) and gibberellin 2-oxidase-like (Paulownia_LG10G000034.1) in PFI were significantly higher than in PF, while auxin response factor, genes involved in auxin signal transduction, were down-regulated in PFI. Upregulation of CK/IAA causes imbalance of hormones, which may lead to PaWB symptoms.

Conclusions

The first integrated analysis that included SMRT sequencing, sequencing on an Illumina platform, and HPLC-MS was performed to investigate changes in transcripts and metabolites in phytoplasma-infected *P. fortunei*. We obtained 5,606 transcripts, 46 lncRNAs, 83 fusion transcripts, 597 TFs, and 107 metabolites. Subsequent analysis of the transcriptome and metabolome found trans-cinnamate 4-monooxygenase, caffeoyl-CoA-O-methyltransferase, ferulic acid, and peroxidase were up-regulated in PFI. These were beneficial to flavonoid and lignin production, which phenylpropane metabolism was enhanced to resist phytoplasma-infection Paulownia. Full-length transcriptome and secondary metabolism analysis revealed a novel regulation mechanism, which will enrich the understanding of Paulownia–phytoplasma interactions. This study will help in better understanding the changes in genes, metabolites, and morphology in Paulownia with PaWB disease.

Methods

Plant materials and MMS treatment

All the *P. fortunei* healthy seedlings (PF) and the corresponding phytoplasma infected seedlings (PFI) used in this study were tissue cultured seedlings, coming from the Institute of Paulownia, Henan Agricultural University, China. The situation of the phytoplasma in the above two preserved seedlings have been identified by Jie (2009) [51]. Before the two seedlings were used in this study, the two seedlings were firstly grown on 1/2 Murashige-Skoog (MS) [52]for 30 days, then the terminal buds of 1.5cm from PF and PFI were transferred into 1/2 MS which contained 0 and 60 mg·L⁻¹ MMS (PF, PFI, PF-60, PFI-60), respectively. Three terminal bud of each sample were cultured in one columnar flask, and 30 bottles were cultivated for each sample. The methods and conditions of these seedlings' cultivation were described by Zhai et al [5]. After 30 days, the terminal buds with 1.5cm length of the four sample were collected, and then frozen in liquid nitrogen and stored at -80°C for RNA extraction. In this study, three biological replicates were performed for the seedling culture.

Total RNA extraction

Total RNAs were extracted from all samples (four groups, each with three samples total 12 groups), with TRIzol reagent (Life Technologies, Dallas, USA). The integrity and purity of RNA samples were conducted using Agilent 2100 bioanalyzer (Agilent Technologies, Santa Clara, CA, USA) and 2000 ultramicro NanoDrop spectrophotometer (Thermo Scientific, Wilmington, DE, USA) according to the manufacturer's instruction. The obtained total RNAs were stored in -80°C for later use.

Illumina Hiseq X-ten library construction and sequencing

The Illumina Hiseq X-ten library was constructed using TruSeq RNA kit (Invitrogen, Carlsbad, CA, USA) and sequenced using Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA, USA) and ABI Step One Plus real-time PCR System (Thermo Scientific, Wilmington, DE, USA).

PacBio Iso-Seq library construction and sequencing

The full-length cDNA was synthesized with the SMARTer™ PCR cDNA Synthesis Kit (Clontech, TAKARA, Japan). The cDNA SMRTbell™ library was constructed with PacBio DNA template Kit (Pacific Biosciences, USA) according to the manufacturer's specifications. To avert loading bias of SMRT sequencing, four different libraries, with cDNA insert size (1-2 kb, 2-3 kb, and 3-6 kb) were prepared, respectively. The concentration and quality of the resultant PCR were assessed Qubit system. CDNA SMRTbell™ library was constructed with PacBio DNA template kit (Pacific Biosciences, USA) according to the manufacturer's instruction.

Alternative splicing

In order to analyze the full-length transcriptome information of Alternative splicing (AS) for the samples, TransDecoder software (v3.0.0) was performed. Reads were mapped to the reference genome of *P. fortunei* to analyze the alternative splicing site by BLAST software (version 2.2.26). And gffcompare software whose purpose was to compare known transcripts with the obtained transcripts by sequencing.

Long non-coding RNA prediction

To ascertain more newly transcripts annotation information, then the Coding Potential Calculator (CPC), Coding-Non-Coding Index (CNCI), Pfam protein domain analysis (Pfam), and Coding Potential Assessment Tool (CPAT) were conducted to predict the coding potential of new transcripts. Subsequently, using EMBOSS software, Open Reading Frame (ORF) prediction was performed on the identified lncRNAs and the amino acid sequence length of more than 100 was filtered out to obtain the final lncRNAs prediction.

Identification of fusion transcripts

Fusion transcripts obtained from RS-IsoSeq analysis were predicted. The methods were referred to Chao et al [30]. In addition, fusion transcripts were validated by Illumina short read of the HiSeqX Ten platform.

Transcript annotation

All the transcripts were mapped to the reference genome, searched against the NCBI non-redundant protein sequence database (Nr) (<http://www.ncbi.nlm.nih.gov/>), the GO with Blast2 GO program, Cluster of Orthologous Groups (COG) database (<http://www.ncbi.nlm.nih.gov/COG>), the Kyoto Encyclopedia of Genes and Genomes (KEGG) databases (<http://www.genome.jp/kegg/>) and Pfam database using BLAST. And GO and COG function classification was carried out. To better understand the biological functions of transcripts related to PaWB, KEGG pathway analysis was performed.

Differential expression analysis of different samples

The expression of genes and transcripts was quantified using RSEM by Maximization tools. Differentially expressed genes and transcripts were analyzed using DESeq. The False Discovery Rate (FDR) was ≤ 0.01 and Fold Change (FC) was ≥ 1.5 were the standard of differentially expressed genes. The FDR was corrected by means of the p-value. According to this standard, differentially expressed transcripts related to PaWB were identified.

Quantitative real-time polymerase chain reaction (qRT-PCR) verification of differentially expressed transcripts

Total RNA was extracted using TRIzol (Sangon, Shanghai, China) for qRT-PCR analysis, repetition of three experiments. According to Chen et al reported design as the primers for PCR amplification (Chen et al. 2005). PCR program was referred to our previous study [53]. 18S rRNA was selected as the internal reference gene. The relative expression is calculated using $2^{-\Delta\Delta C_t}$.

Isoform identification and real-time polymerase chain reaction (RT-PCR)

To distinguish different alternative splicing events, PacBio-seq isoform sequences were designed Isoform-specific primers (Table S3). Total RNA was extracted from the four samples (PF, PFI, PF-60, and PFI-60) using TRIzol (Sangon, Shanghai, China) for qRT-PCR analysis, repetition of three experiments. According to Chen et al reported design as the primers for PCR amplification[1]. PCR program was referred to our previous study [53]. 18S rRNA was selected as the internal reference gene. The relative expression is calculated using $2^{-\Delta\Delta C_t}$.

Metabolite extraction and analysis by high performance liquid chromatography-mass spectrometry (HPLC-MS)

The extraction method of metabolites of plant leaf were referred to Cao [29], the metabolite was extracted from the samples using LC-ESI-MS/MS (HPLC, Shim-pack UPLC SHIMADZU CBM20A system). Applied Biosystems 4000 Q TRAP system was performed, for data analysis. And the dates were processed with software Analyst 1.6.1 (AB SCIEX, CA, USA).

Screening differential metabolites related to PaWB

Principal Component Analysis (PCA) and Partial Least Squares-Discriminant Analysis (PLS-DA) were used to analyze the difference metabolites between different samples. PCA was used to determine the separation trend among sample groups. Different varieties or tissue metabolites were screened firstly by multivariate analysis with PLS-DA VIP parameter values of the model. Then numerical fold changes were used to further screening. The difference multiple is the ratio of the metabolite content in the treatment group divided by the metabolite content in the control group. The resulting difference metabolites should meet the following conditions: fold change is greater than or equal to 2 or less than 0.5, and the VIP is greater than or equal to 1.

Correlation analysis between metabolites and transcriptome

In this study, pearson correlation coefficients (> 0.8) were designed for metabolome and transcriptome data integration. A core program from R was used to determine the correlation between the number of enzymes corresponding to the enzyme in KEGG pathway and the screening criteria for the Correlation results.

Abbreviations

P. fortunei: *Paulownia fortunei*; PaWB: Paulownia withes'-broom; MMS: Methyl methane sulfonate; SMRT: Single molecule real time; HPLC-MS: High-performance liquid chromatography-mass spectrometry; AS: Alternative splicing; PF: *Paulownia fortunei* of healthy seedlings; PFI: *Paulownia fortunei* with phytoplasma infection; PF-60: Healthy seedlings treated with 60 mg•L⁻¹ methyl methane sulfonate; PFI-60: Phytoplasma-infected seedlings treated with 60 mg mg•L⁻¹ methyl methane sulfonate; ROI: Read of insert; QRT-PCR: Quantitative real-time polymerase chain reaction; RT-PCR: Real-time polymerase chain reaction; IR: Intron retention; LncRNAs: Long non-coding RNAs; CK: Cytokinin; FPKMs: Fragments per kilobase of transcript per million fragments mapped; CPC: Coding potential calculator; CNCI: Coding-non-coding index; Pfam: Pfam protein domain analysis; CPAT: Coding potential assessment tool; ORF: Open reading frame; Nr: Non-redundant protein sequence database; GO: Gene ontology; COG: Cluster of orthologous groups; KEGG: Kyoto encyclopedia of genes and genomes; FDR: False discovery rate; FC: Fold change; PCA: Principal component analysis; PLS-DA: Partial least squares-discriminant analysis; WUS: WUSCHEL; PfwUS: Paulownia_LG5G000525 encoding protein WUS-like; TFs: Transcription factors; WGCNA: Weighted correlation network analysis; QC: Quality control; SPL9: Squamosa promoter binding protein-like 9; *Pfwus*: Gene Paulownia_LG5G000525.1; ARR: Two-component ARABIDOPSIS RESPONSE REGULATOR

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and material

The Illumina platforms sequencing data reported in this study are available in the Sequence Read Archive (<http://www.ncbi.nlm.nih.gov/sra>) under Accession No. SRP192794. SMRT sequencing data reported in this study are available in the Sequence Read Archive (<http://www.ncbi.nlm.nih.gov/sra>) under Accession No. SRP193021.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by the National key research and development program (2017YFD060050604), the Distinguished Talents Foundation of Henan Province of China (174200510001), and the Natural Science Foundation of Henan Province of China (162300410158). The funding bodies had no roles in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

Authors' contributions

GF and XZ² designed the experiments. XZ¹ wrote the paper. WZ, YC and ZZ performed the data. XC analyzed the data. GF, and XC revised the paper. All authors read and approved the final manuscript.

Acknowledgment

We thank Margaret Biswas, PhD, from Liwen Bianji, Edanz Group China (www.liwenbianji.cn/ac), for editing the English text of a draft of this manuscript.

References

1. Fan G, Cao X, Niu S, Deng M, Zhao Z, Dong Y. Transcriptome, microRNA, and degradome analyses of the gene expression of Paulownia with phytoplasma. BMC GENOMICS 2015, 16:896.
2. Doumett S, Lamperi L, Checchini L, Azzarello E, Mugnai S, Mancuso S et al. Heavy metal distribution between contaminated soil and Paulownia tomentosa, in a pilot-scale assisted phytoremediation study: influence of different complexing agents. CHEMOSPHERE 2008, 72(10):1481-1490.
3. Du T, Wang Y, Hu Q, Chen J, Liu S, Huang W et al. Transgenic Paulownia expressing shiva-1 gene has increased resistance to Paulownia witches' broom disease. J INTEGR PLANT BIOL 2005, 47(12):1500-1506.
4. Lee IM, Davis RE, Gundersen-Rindal DE. Phytoplasma: phytopathogenic mollicutes. Annu Rev Microbiol 2000, 54:221-255.
5. Zhai X, Cao X, Fan G. Growth of paulownia witches' broom seedlings treated with methylmethane sulphonate and SSR analysis. Sci Silv Sin 2010, 46(12):176-181.
6. Namba S. Phytoplasmas: A century of pioneering research. J PLANT PATHOL 2011, 77(6):345-349.
7. Zhai X, Wang G, Bi H, Li P. The relationship between Paulownia witches' broom and phenol and amino acid change in leaves. Henan Science 2000, 18(3):277-279.
8. Tian G, Huang Q, Yuan Q, Zhang X. Correlation of metabolic changes of infected paulownia tissue culture with PWB-MLO pathogenic mechanism. Science in China 1995(8):934-943.
9. Wang R. Effects of hormones on the morphogenesis of witches broom of Paulownia spp. Sci Silv Sin 1981, 17(3):281-286.
10. Liu R, Dong Y, Fan G, Zhao Z, Deng M, Cao X et al. Discovery of genes related to witches broom disease in Paulownia tomentosa × Paulownia fortunei by a de novo assembled transcriptome. PLOS ONE 2013, 8(11):e80238.
11. Fan G, Cao X, Niu S, Deng M, Zhao Z, Dong Y. Transcriptome, microRNA, and degradome analyses of the gene expression of Paulownia with phytoplasma. BMC GENOMICS 2015, 16(1):1-15.
12. Niu S, Fan G, Deng M, Zhao Z, Xu E, Cao L. Discovery of microRNAs and transcript targets related to witches' broom disease in Paulownia fortunei by high-throughput sequencing and degradome approach. MGG 2016, 291(1):181-191.

13. Cao X, Fan G, Dong Y, Zhao Z, Deng M, Wang Z et al. Proteome profiling of Paulownia seedlings infected with phytoplasma. *Front Plant Sci* 2017, 8:342.
14. Fan G, Wang L, Dong Y, Zhao Z, Deng M, Niu S et al. Genome of Paulownia (*Paulownia fortunei*) illuminates the related transcripts, miRNA and proteins for salt resistance. *Scientific reports* 2017, 7(1):1285.
15. Fan G, Cao Y, Wang Z. Regulation of long noncoding RNAs responsive to phytoplasma infection in *Paulownia tomentosa*. *Int J Genomics* 2018, 22:1-16.
16. Wang Z, Li B, Li Y, Zhai X, Dong Y, Deng M et al. Identification and characterization of long noncoding RNA in *Paulownia tomentosa* treated with methyl methane sulfonate. *PHYSIOL MOL BIOL PLA* 2018, 24(2):325-334.
17. Dixon RA, Strack D. Phytochemistry meets genome analysis, and beyond. *PHYTOCHEMISTRY* 2003, 62(6):815-816.
18. Ward JL, Baker JM, Beale MH. Recent applications of NMR spectroscopy in plant metabolomics. *FEBS J* 2010, 274(5):1126-1131.
19. Xu Q, Zhu J, Zhao S, Hou Y, Li F, Tai Y et al. Transcriptome profiling using single-molecule direct RNA sequencing approach for in-depth understanding of genes in secondary metabolism pathways of *Camellia sinensis*. *Front Plant Sci* 2017, 8:1205.
20. Workman R, Myrka A, Wong G, Tseng E, Welch JK, Timp W. Single-molecule, full-length transcript sequencing provides insight into the extreme metabolism of the ruby-throated hummingbird *Archilochus colubris*. *GIGASCIENCE* 2018, 7(3):1-12.
21. Wang B, Tseng E, Regulski M, Clark TA, Hon T, Jiao Y et al. Unveiling the complexity of the maize transcriptome by single-molecule long-read sequencing. *NAT COMMUN* 2016, 7:11708.
22. Ponting CP, Oliver PL, Reik W. Evolution and functions of long noncoding RNAs. *CELL* 2009, 136(4):629-641.
23. Wang J, Tian C, Zhang C, Shi B, Cao X, Zhang T et al. Cytokinin signaling activates WUSCHEL expression during axillary meristem initiation. *PLANT CELL* 2017, 29(6):1373-1387.
24. Nacu S, Yuan W, Kan Z, Bhatt D, Rivers CS, Stinson J et al. Deep RNA sequencing analysis of readthrough gene fusions in human prostate adenocarcinoma and reference samples. *BMC MED GENOMICS* 2011, 4(1):11.
25. Weckselblatt, Brooke, Rudd, Katharine M. Human structural variation: mechanisms of chromosome rearrangements. *TRENDS GENET* 2015, 31(10):587-599.
26. Sun Z, Su C, Yun J, Jiang Q, Wang L, Wang Y et al. Genetic improvement of the shoot architecture and yield in soybean plants via the manipulation of GmmiR156b. *Plant biotechnology journal* 2018:1-13.
27. Schoof H, Lenhard M, Haecker A, Mayer KF, Jürgens G, Laux T. The stem cell population of *Arabidopsis* shoot meristems is maintained by a regulatory loop between the CLAVATA and WUSCHEL genes. *CELL* 2000, 100(6):635-644.

28. Leibfried A, To J, Busch W, Stehling S, Kehle A, Demar M et al. WUSCHEL controls meristem function by direct regulation of cytokinin-inducible response regulators. *Nature* 2005, 438(7071):1172-1175.
29. Cao Y, Zhai X, Deng M, Zhao Z, Fan G, University HA. Relationship between metabolites variation and *Paulownia witches' broom*. *Sci Silv Sin* 2017, 56(3):85-93.
30. Chao Q, Gao ZF, Zhang D, Zhao BG, Dong FQ, Fu CX et al. The developmental dynamics of the *Populus* stem transcriptome. *Plant biotechnology journal* 2018, 196:1-14.
31. Ocwieja KE, Scott SM, Rithun M, Rebecca CA, Patricia D, Michael B et al. Dynamic regulation of HIV-1 mRNA populations analyzed by single-molecule enrichment and long-read sequencing. *NUCLEIC ACIDS RES* 2012, 40(20):10345-10355.
32. Zhang W, Ciclitira P, Messing J. PacBio sequencing of gene families - a case study with wheat gluten genes. *GENE* 2014, 533(2):541-546.
33. Abdelghany SE, Hamilton M, Jacobi JL, Ngam P, Devitt N, Schilkey F et al. A survey of the sorghum transcriptome using single-molecule long reads. *NAT COMMUN* 2016, 7:11706.
34. Wang M, Wang P, Liang F, Ye Z, Li J, Shen C et al. A global survey of alternative splicing in allopolyploid cotton: landscape, complexity and regulation. *The New phytologist* 2018, 217(1):163-178.
35. Wilson ID, Robert P, Jennifer G, Hilary M, Rebecca W, Lenz EM. HPLC-MS-based methods for the study of metabonomics. *J Chromatogr B Analyt Technol Biomed Life Sci* 2005, 817(1):67-76.
36. Liang Y, Hu X, Zhang Y, Liu X. Progress on physiological function research of plant peroxidase. *Journal of Inner Mongolia Agricultural University* 2003, 24:100-104.
37. Kim JB. Pathogen, Insect and Weed Control Effects of Secondary Metabolites from Plants. *J KOREAN SOC APPL BI* 2005, 48:1-15.
38. Passardi F, Penel C, Dunand C. Performing the paradoxical: how plant peroxidases modify the cell wall. *TRENDS PLANT SCI* 2004, 9(11):534-540.
39. Zhouen Z, Side Y, Weizhen L, Wenfeng W, Yizun J, Nianyun L. Mechanism of reaction of nitrogen dioxide radical with hydroxycinnamic acid derivatives: a pulse radiolysis study. *FREE RADICAL RES* 1998, 29(1):13-16.
40. Mahjoub A, Hernould M, Joubès J, Decendit A, Mars M, Barrieu F et al. Overexpression of a grapevine R2R3-MYB factor in tomato affects vegetative development, flower morphology and flavonoid and terpenoid metabolism. *PLANT PHYSIOL BIOCH* 2009, 47(7):551-561.
41. Liu J, Osbourn A, Ma P. MYB transcription factors as regulators of phenylpropanoid metabolism in plants. *MOL PLANT* 2015, 8(5):689-708.
42. Margaria P, Ferrandino A, Caciagli P, Kedrina O, Schubert A, Palmano S. Metabolic and transcript analysis of the flavonoid pathway in diseased and recovered Nebbiolo and Barbera grapevines (*Vitis vinifera* L.) following infection by *Flavescence dorée* phytoplasma. *PLANT CELL ENVIRON* 2014, 37(9):2183-2200.

43. Li H, Zhou Y, Zhang Z. Network analysis reveals a common host-pathogen interaction pattern in Arabidopsis immune responses. *Front Plant Sci* 2017, 8:893.

44. Finlayson SA. Arabidopsis thosinte branched1-like1 regulatesaxillary bud outgrowth and is homologous to monocot teosinte branched1. *PLANT CELL PHYSIOL* 2007, 48(5):667-677.

45. Bell EM, Lin WC, Husbands AY, Yu L, Jaganatha V, Jablonska B et al. Arabidopsis lateral organ boundaries negatively regulates brassinosteroid accumulation to limit growth in organ boundaries. *PNAS* 2012, 109(51):21146-21151.

46. Tanaka W, Ohmori Y, Ushijima T, Matsusaka H, Matsushita T, Kumamaru T et al. Axillary meristem formation in Rice requires the WUSCHEL ortholog tillers absent1. *PLANT CELL* 2015, 27(4):1173-1184.

47. Lu Z, Shao G, Xiong J, Jiao Y, Wang J, Liu G et al. MONOCULM 3, an ortholog of WUSCHEL in rice, is required for tiller bud formation. *J GENET GENOMICS* 2015, 42(2):71-78.

48. Kiba T, Yamada HS, Kato T, Tabata S, Yamashino T, Mizuno T. The type-A response regulator, ARR15, acts as a negative regulator in the cytokinin-mediated signal transduction in Arabidopsis thaliana. *PLANT CELL PHYSIOL* 2003, 44(8):868-874.

49. To JPC, Georg H, Ferreira FJ, Jean D, Mason MG, G Eric S et al. Type-A Arabidopsis response regulators are partially redundant negative regulators of cytokinin signaling. *PLANT CELL* 2004, 16(3):658-671.

50. Andrea L, To JPC, Wolfgang B, Sandra S, Andreas K, Monika D et al. WUSCHEL controls meristem function by direct regulation of cytokinin-inducible response regulators. *Nature* 2005, 438(7071):1172-1175.

51. Jie D, Zhai X, Nie L. Phytoplasma detection and identification techniques. *Henan Forestry Science and Technology* 2009, 29(1):39-41.

52. Murashige T, Skoog F. A Revised Medium for Rapid Growth and Bio Assays with Tobacco Tissue Cultures. *Physiologia Plantarum* 1962, 15(3):473-497.

53. Fan G, Niu S, Zhao Z, Deng M, Xu E, Wang Y et al. Identification of microRNAs and their targets in Paulownia fortunei plants free from phytoplasma pathogen after methyl methane sulfonate treatment. *BIOCHIMIE* 2016, 127:271-280.

Tables

Table 1 Statistics of the full-length reads in *P. fortunei*

Sample Name	cDNA size	Reads of Insert	Number of 5'prime reads	Number of 3'prime reads	Number of poly-A reads	Number of non-full-length reads	Number of full-length reads	Number of full-length non-chimeric reads	Average full-length non-chimeric read length
PF	1-2K	121,148	59,456	68,270	65,972	53,809	48,642	48,404	1,333
	2-3K	91,707	55,704	59,695	59,014	39,206	47,091	47,075	2,436
	3-6K	63,367	34,989	37,960	37,667	30,498	29,250	29,233	3,614
PFI	1-2K	142,904	78,979	88,077	85,380	55,971	66,354	66,123	1,243
	2-3K	121,908	74,971	80,246	79,523	50,142	64,022	63,989	2,346
	3-6K	67,591	44,409	46,446	46,124	28,334	37,654	37,642	3,516
PF-60	1-2K	90,554	48,370	53,927	52,462	37,100	40,694	40,424	2,371
	2-3K	113,908	61,163	67,691	66,813	52,833	51,425	51,400	3,495
	3-6K	107,025	58,692	63,391	63,322	53,677	48,970	48,883	1,284
PFI-60	1-2K	140,811	76,784	85,874	83,427	56,020	64,432	64,240	2,254
	2-3K	126,096	75,841	81,035	80,602	52,523	64,407	64,387	3,292
	3-6K	117,450	71,585	75,380	75,048	53,361	60,520	60,495	2,777

Figures

Figure 1

SMRT sequencing quality testing. (a) The results of full-length transcriptome integrity assessment results. (b) Box-plots show the dispersion degree of transcriptome expression level distribution. (c) The correlation analysis of the samples. (d) The qRT-PCR validation of transcripts, Standard error of the mean for three biological replicates is represented by the error bars. (1): Protein phosphatase 2C; (2): NAC domain-containing protein 100-like; (3): UDP-galactose/UDP-glucose transporter 4-like; (4): Probable serine/threonine-protein kinase Cx32, chloroplastic; (5): Disease resistance protein; (6): Haloacid dehalogenase-like hydrolase domain-containing protein 3.

Figure 2

Alternative splicing events and different isoforms from Iso-seq. (a) Statistics of different alternative splicing type in *P. fortunei*. (b) Distribution of the number of isoforms per gene. The average numbers of isoforms for all genes in reference annotation and PacBio data are shown in red and green, respectively. (c) A pie chart showing the percentage of PacBio Iso-Seq isoforms that are the same as existing gene models, novel isoforms of known genes and novel isoforms of novel genes. (d) RT-PCR validation of AS isoforms for six genes. Gel bands in each Figure show DNA makers and PCR results in six samples. Blue boxes show exons and lines with arrows show introns. PCR primers are shown on the first isoform of each gene. The length of each full-length isoform is shown after the transcript structure.

Figure 3

Statistics of the lncRNA in PaWB. (a) According to their positions in the RefGen_v3 annotations. (b) The GO functional classification and functional annotation of differentially expressed target genes.

Figure 4

CIRCOS visualization of genomic and transcriptomic features. (a) Paulownia chromosomes. (b) Gene density of four samples (PF, PFI, PF-60, and PFI-60). (c) Transcript density of four samples (PF, PFI, PF-60, and PFI-60). (d) Long non-coding RNA (lncRNA) distribution. (e) Fusion transcript distribution: intra-chromosome (purple); inter-chromosome (yellow).

Figure 5

K-means clustering and enrichment analysis of transcripts differentially expressed in PF, PFI, PF-60, and PFI-60.

(a) K-means clustering showing the transcriptome expression profiles. Ten clusters were identified based on expression levels in four samples (PF, PFI, PF-60, and PFI-60). (b) Gene Ontology enrichment among the ten clusters. White, not significant, yellow to red, significant enrichment.

Figure 6

Weighted Gene Co-expression Network Analysis (WGCNA) of transcription factors. (a) Cluster dendrogram of transcription factors based on expression levels in the four developmental zones (PF, PFI, PF-60, and PFI-60).

Each branch represents a gene and each color below represents a gene co-expression module. Dynamic tree cut indicates the modules divided based on the gene clustering results. Merged dynamic indicates the modules divided by combining modules with similar expression patterns. (b) Distribution of transcription factor families in KEGG pathways.

Figure 7

Metabolites. (a). Clustering heat map of different sample metabolites of Paulownia, red in the heat map indicates higher substance content, blue indicates lower substance content. (b)TIC overlap Figure of QC sample. (c) PCA analysis of metabolites. (d) The network of the difference genes and differential metabolites. The metabolites in the Figure are marked in red, and the correlation between them is indicated by a straight line.

Figure 8

The phenylpropanoid biosynthesis pathway. (a) The network of the difference genes and differential metabolites. The metabolites in the Figure are marked in red, and the correlation between them is indicated by a straight line. (b) The phenylpropanoid biosynthesis pathway. The metabolites and transcripts up-regulated are marked in red.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [supplement1.docx](#)
- [supplement2.xlsx](#)
- [supplement3.docx](#)
- [supplement4.docx](#)

- [supplement5.xlsx](#)
- [supplement5.docx](#)
- [supplement7.doc](#)
- [supplement8.xlsx](#)
- [supplement9.tif](#)
- [supplement10.tif](#)
- [supplement11.tif](#)