

Incomplete Genome Doubling Enables to Consistently Enhance Plant Growth for Maximum Biomass Production by Altering Multiple Transcript Co-Expression Networks in Potato

Kanglu Zhao

Inner Mongolia University

Nengzhou Jin

Inner Mongolia University

Meysam Madadi

Huazhong Agriculture University

Youmei Wang

Huazhong Agriculture University

Lei Wu

Inner Mongolia University

Zhijun Xu

Inner Mongolia University

Jinxuan Wang

Inner Mongolia University

Jing Dong

Inner Mongolia University

Shang-wen Tang

Huazhong Agriculture University

Yanting Wang

Huazhong Agriculture University

Liangcai Peng

Huazhong Agriculture University

Zhiyong Xiong (✉ xiongzy2003@aliyun.com)

Inner Mongolia University <https://orcid.org/0000-0003-3073-4649>

Research Article

Keywords: potato, polyploid, cytochimera, biomass, co-expression network

Posted Date: June 9th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-578375/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Theoretical and Applied Genetics on November 3rd, 2021. See the published version at <https://doi.org/10.1007/s00122-021-03976-y>.

Abstract

Polyploidization is an important approach in crop breeding for agronomic trait improvement, especially for biomass production. Cytochimera contains two or more mixed cells with different levels of ploidy, which is considered a failure in whole genome duplication. Using colchicine treatment with diploid (Dip) potato (*Solanum chacoense*) seedlings, this study generated tetraploid (Tet) and cytochimera (Cyt) lines, which respectively contained complete and partial cells with genome duplication. Compared to the Dip potato, we observed remarkably enhanced plant growth and biomass yields in Tet and Cyt lines. Notably, the Cyt potato straw, which was generated from incomplete genome doubling, was of significantly higher biomass yield than that of the Tet with a distinctively altered cell wall composition. Meanwhile, we observed that one layer of the tetraploid cells (about 30 %) in Cyt plants was sufficient to trigger a gene expression pattern similar to that of Tet, suggesting that the biomass dominance of Cyt may be related to the proportion of different ploidy cells. Further genome-wide analyses of co-expression networks indicated that down-regulation (against Dip) of spliceosomal-related transcripts might lead to differential alternative splicing for specifically improved agronomic traits such as plant growth, biomass yield and lignocellulose composition in Tet and Cyt plants. In addition, this work examined that the genome of Cyt line was relatively stable after years of asexual reproduction. Hence, this study has demonstrated that incomplete genome doubling is a promising strategy to maximize biomass production in potato and beyond.

Introduction

Polyploidization is a basic feature of plant evolution and diversification via genome duplication (Wendel 2000), and polyploid plants often exhibit increment of organs, buffering of deleterious mutations, increased heterozygosity, and growth vigor (Cai et al. 2007; Corneillie et al. 2019; Shelley et al. 2018). Regarding such features, polyploidization thus provides an effective non-transgenic breeding approach to increase grain and biomass yields in crops (Cai et al. 2007; Corneillie et al. 2019; Shelley et al. 2018). For instance, the application of polyploidization has shown remarkably increased biomass in bioenergy *Miscanthus* (Chae et al. 2013) and willow crops (Serapiglia et al. 2015). Notably, the tetraploid *Arabidopsis* plants have almost doubled biomass yields than those of the diploid ones, but the hexaploid and octoploid *Arabidopsis* plants do not show any significantly higher biomass yields than those of the diploid ones, raising a question if the maximum biomass yield could only be achieved in the tetraploid crops (Corneillie et al. 2019). In addition, the polyploidization causes a constant reduction of lignin content, leading to significantly improved lignocellulose recalcitrance, which is beneficial to bioethanol production (Corneillie et al. 2019; Madadi et al. 2021).

In principle, there are two different polyploid types in plants: allopolyploids derived from interspecific hybridization followed by genome doubling, and autopolyploids originated from a single species (Comai 2005; Otto 2007). Allopolyploidy can be confounded by the entanglement of both whole genome duplication (WGD) and hybridization (Zhang et al. 2015). For instance, investigations in maize indicated that gene expression is altered less by WGD than by hybridization (Auger et al. 2005). In contrast,

synthesized autopolyploids represent ideal materials to investigate the consequences of genome duplication, especially to ascertain whether there are molecular and functional rules about increasing biomass following polyploidization events (Fasano et al. 2016). However, the consequences of chromosome doubling have been widely studied in allopolyploids, whereas more limited data are available for autopolyploid plants (Parisod et al. 2010). Extensive studies of genetic and epigenetic changes associated with polyploidization have been demonstrated that chromosomal rearrangements, sterility, aneuploidy, gene-expression aberrations, and DNA methylation changes are common in newly induced polyploids (neopolyploids) (Adams and Wendel 2005; Comai 2005; Ha et al. 2009; Leitch and Leitch 2008; Song et al. 1995; Xiong et al. 2011). However, few investigations are explored to examine the associations between growth vigor and gene expression in neopolyploids response to WGD.

A plant cytochimera is defined as an individual composed of two or more groups of cells that are genetically different for ploidy. In the case of cultured tissues, which are often induced spontaneously by somaclonal variations or artificially by mutagen treatment that lead to "cytochimera" (Frank and Chitwood 2016). For instance, polyploids can be induced by colchicine treatment of lateral buds, while cytochimeras may arise in the case as a result of incomplete penetration of colchicine (Hashimoto-Freitas and Nassar 2013). Sometimes cytochimera presence is more valuable than that of the normal type, especially in an ornamental plant, but in most cases, it is considered to negatively affect breeding performance (Fujishige et al. 1996). Since the characteristics of a mixture of different ploidy cells, unstable heredity, and usually without desirable traits, cytochimeras are deliberately avoided and abandoned during the plant polyploid breeding.

Because there is a variety of wild or domesticated diploid, tetraploid and hexaploid materials (Carputo et al. 2003), its special tolerance on ploidy manipulation (Ortiz and Ehlenfeldt 1992), and the strict tetraploid inheritance of cultivated potato (Carputo et al. 2003), the potato is a good model plant for autopolyploid research. Cultivated autopolyploid potatoes are often more vigorous and larger in size than their diploid relatives, indicating a fitness advantage associated with higher ploidy (Stupar et al. 2007). However, the diploid *Solanum phureja* is most vigorous and generates the greatest biomass than its derived monopleid and synthetic autopolyploid (Stupar et al. 2007). Synthetic autotetraploid of *S. commersonii* and *S. bulbocastanum* have no growth vigor compared with their diploid progenitors (Fasano et al. 2016). In this study, therefore, we generated tetraploid (Tet) and cytochimera (Cyt) potato lines, and observed consistently enhanced plant growth in both polyploidy lines. We then determined distinctively altered lignocellulose composition in the Cyt straw. Finally, this study attempted to sort out why the Cyt potato line was of the highest biomass yield by performing a genome-wide analysis of transcript co-expression networks, providing a promising strategy to maximize biomass yields in crop breeding.

Materials And Methods

Plant growth and collection

Seeds of wild diploid potato (*Solanum chacoense*, PI 500042) were obtained from the US potato gene bank. Sterilized seeds were germinated to form tissue culture seedlings on MS basal medium containing 3 % sucrose and 0.6 % agar without exogenous growth regulators (pH 5.8). For all experiments, seedlings were propagated from one original tissue culture plant through a cutting node with an axillary bud and incubated in vitro under the fixed conditions (16 h light/8 h dark, 24 °C). The tetraploid (Tet-1, 2) and cytochimera (Cyt-1, 2) lines were generated by colchicine treatment with diploid (Dip) plants. To produce minitubers, about fifty well-grown propagated seedlings of each ploidy were cultivated in the greenhouse, then the minitubers were grown in the field for ploidy identification and phenotype determination. The harvested tubers were grown in the field to repeat the experiment and to collect stems and leaves for paraffin section, cell wall characterization, and RNA-seq.

Ploidy plant identification

A small section of node (approximately 1 cm length) with an axillary bud from propagated diploid was cultured onto the MS medium supplied with various concentration of colchicine (0.025 %, 0.05 %, 0.075 %, and 0.1 %) for 10 days, and then transferred to the medium without colchicine. The 0.05 % colchicine was finally selected for polyploidy induction experiments to screen and determine the somatic ploidy level. A small part of the leaf (2 mm²) was fixed in Carnoy's fluid and digested with 1% pectolyase Y-23 and 2% cellulase R-10, and the ploidy was then identified using fluorescence in situ hybridization (FISH) method with 5S rDNA as probes on the somatic nucleus. Thus, the diploid has one pair of 5S rDNA locus located on chromosome one, and two signals of 5S rDNA on the somatic nucleus, while tetraploid by doubling the genome has four hybrid signals (Dong et al. 2000), but cytochimera cells contain two or four signals. To detect the percentage of tetraploid cells in the cytochimera plants from different lines, somatic nuclei from various organs of root, stem, leaf, and flower at the flowering stage were prepared for 5S rDNA FISH. More than 500 nuclei were randomly photographed and the percentage of tetraploid cells to total cells was calculated. To determine the chromosome integrality of field-grown plants, multi-color FISH using 5S rDNA, 45S rDNA, and chromosomal specific BAC clone BAC079E02 as probes was performed as previously described (Xiong et al. 2021; Xiong et al. 2011).

Phenotype measurement

The well-grown plants of each line were randomly selected to determine the height (cm), number of main stems, leaf coverage area (cm²), pollen activity (%), yield of tubers (g), fresh and dry weights (g) at the young, flowering, and mature stages using the standard methods. The fresh straws were weighed and then dried at 50 °C for 3-5 days as dry matter. Three biological replications were carried out for each line and significance analysis was performed by the LSD test at $P < 0.05$.

Cell wall polymer extraction and assay

Plant cell wall fractionation was performed as previously described (Peng et al. 2000). The well-mixed biomass powders were extracted with potassium phosphate buffer (pH 7.0), chloroform-methanol (1:1, v/v), DMSO–water (9:1, v/v) and ammonium oxalate 0.5% (w/v) to obtain lipid, starch, and pectin,

respectively. The remaining crude cell walls were extracted with 4 M KOH containing 1.0 mg/mL sodium borohydride for 1 h at 25 °C and the combined supernatants were used as the hemicelluloses fraction. The hexose of the remaining final pellet was determined as cellulose level by treatment with H₂SO₄ (67%, v/v) for 1 h at 25 °C using the anthrone/H₂SO₄ method (Fry et al. 1988). Total hemicellulose levels were calculated by measuring all hexoses and pentoses in the hemicellulose fraction and pentoses in the final pellet using the orcinol/HCL method (Dische et al. 1962). Total pectin was calculated by measuring hexoses, pentoses, and uronic acids of the pectin fraction. Total lignin was measured according to the Laboratory Analytical Procedure of the National Renewable Energy Laboratory (Sluiter et al. 2008). All experimental analyses were performed in independent triplicates.

RNA sequencing

The well-grown plants in the field from Dip, Tet-1, and Cyt-1 lines were collected for RNA sequencing. Total RNA was extracted from fresh stems and leaves using an RNeasy Plant Mini Kit (Qiagen). The individual cDNA was synthesized using the random hexamers as primers and mRNA templates. The resultant products were connected with adapters, followed by size selection and PCR amplification. Then, the constructed library was analyzed by an Illumina HiSeq X Ten sequencing platform (Novogene, Beijing, China).

Co-expression network analysis and functional annotation

The potato reference genome and annotation files were downloaded from Ensembl Plants, and BioMart converted its different ID (<http://plants.ensembl.org/index.html>). Adaptor sequences and low-quality reads were initially filtered, and the remaining ones were called clean reads. Then, FPKM (fragments per kilobase of exon model per million mapped fragments) expression matrix was obtained through the transcript-level analysis with HISAT, StringTie, and Ballgown (Pertea et al. 2016).

Retaining transcripts with FPKM mean value ≥ 2 of all samples, and then $\log_2(\text{FPKM}+1)$ normalized values were used to generate the co-expression networks by the WGCNA package in R (Langfelder and Horvath 2008). The modules were obtained using the automatic network construction function `blockwiseModules` with default settings, except that `TOMType` was signed, the power, `minModuleSize`, and `mergeCutHeight` were 27, 30, and 0.25, relatively. To visualize the expression profiles of the modules, the eigengene (first principal component) for each module was plotted using `ggplot2` in R. To relate the trait measurements with the network, the module eigengenes (MEs) were correlated with the trait data. To identify hub transcripts within the modules, the module membership (MM) for each transcript was calculated based on the Pearson correlation between the expression level and the ME. Transcripts within the module with the highest MM are highly connected within that module. To associate individual transcripts with the traits, we calculated gene significance (GS) as described in the WGCNA package. Then, networks were visualized using Cytoscape 3.7.1 (<https://cytoscape.org/>).

Both annotations of GO (<http://geneontology.org/>) and KEGG (<https://www.kegg.jp/>) were carried out to identify functional transcripts. Subsequently, functional enrichment of GO and KEGG pathway were

analyzed using clusterProfiler package in R, respectively (Yu et al. 2012). The terms were considered to be significantly enriched if FDR (false discovery rate) ≤ 0.05 .

Results

Distinctively increased biomass yields in polyploid potato straws

Based on a colchicine treatment with the diploid (Dip) potato seedlings, this study established efficient methods to screen and identify polyploid plants by counting the fluorescents in situ hybridization (FISH) signals from 5S rDNA probes on the somatic nucleus. Because two characteristic signals of 5S rDNA on the somatic nucleus accountable for Dip plants, we observed four hybrid signals in root, stem, leaf, and flower organs of the tetraploid (Tet) plant at the flowering stage (Figure S1), indicating that the Tet plants were of doubled genomes. Notably, this study also identified the cytochimera (Cyt) potato plants that mixed diploid and tetraploid cells (Figure S1). In detail, two Cyt lines (Cyt-1, Cyt-2) were examined with 24.2 %-41.8 % tetraploid cells (% total cells) in four major organs of potato plants (Table S1), indicating that the Cyt plants were partial duplicated and contained mostly diploid cells. In addition, the ploidy level and the chromosome integrality of Dip, Tet-1, and Cyt-1 from field-grown plants were further verified by FISH using three different chromosomal specific markers located separately on chromosomes 1, 2, and 3 (Fig. 1a). We found that the Tet-1 plant had a complete 48 chromosomes including two pairs of chromosomes 1, 2, and 3 each, while the Cyt-1 plant contained both diploid cells with parental Dip karyotype and tetraploid cells with Tet-1 like karyotype (Fig. 1a).

Among the polyploid potato lines generated in this study, we selected two elite lines of each polyploid type (Tet-1, 2; Cyt-1, 2) to compare their distinct plant growth and biomass yields with the Dip plants (Fig. 1b). In the two-years field experiments, two Tet lines exhibited much-enhanced plant growth and development than those of the Dip plants including main stems (Table S2) and leaf coverage area (Table S3), consistent with the previous findings that genome duplication could cause a remarkable increase of plant growth in tetraploid plants (Corneillie et al. 2019). Hence, compared to the Dip plants, the Tet plants were of consistently higher biomass yields in three major growth stages, with the biomass up to 6.5-fold for fresh weight and 4.2-fold for dry matter at the mature stage (Table 1). Notably, this study examined that the Cyt plants could consistently show more enhanced plant growth and development than those of the Dip plants, with the significantly increased fresh and dry biomass yields by 9.2-fold and 5.4-fold at the mature stage (Table 1). Thus, despite the tetraploid plants were of the highest biomass yields among the polyploid plants examined in the previous studies (Corneillie et al. 2019; Sattler et al. 2016), this study indicates that the cytochimera plants should be optimal to produce maximum biomass in potato plants. We also examined the tuber yields but find a significant decrease in the Cyt and Tet plants comparing to the Dip plants (Table S4). This data indicated that the mixture of diploid and tetraploid cells in Cyt potato crop may be specific for increasing biomass rather than for tuber production.

Table 1

Fresh and dry weight (g/per plant) of the potato diploid (Dip), tetraploid (Tet), and cytochimera (Cyt) samples at three growth stages.

Samples	Fresh/Dry weight	Young	Flowering	Mature
Dip	Fresh weight	24.3 ± 1.5 ^c	72.0 ± 3.5 ^c	153.3 ± 19.4 ^c
	Dry weight	7.6 ± 1.5 ^c	27.3 ± 2.1 ^c	51.6 ± 10.4 ^c
Tet-1	Fresh weight	159.3 ± 3.0 ^b	196.6 ± 6.6 ^b	1011.3 ± 77.1 ^b
	Dry weight	47.3 ± 1.5 ^b	61.0 ± 3.6 ^b	227.3 ± 6.8 ^b
Tet-2	Fresh weight	115.6 ± 3.5 ^b	164.3 ± 9.7 ^b	988.6 ± 37.7 ^b
	Dry weight	39.0 ± 1.0 ^b	56.0 ± 4.0 ^b	202.3 ± 5.8 ^b
Cyt-1	Fresh weight	279.0 ± 1.5 ^a	341.3 ± 15.0 ^a	1391.0 ± 62.0 ^a
	Dry weight	62.3 ± 1.5 ^a	99.0 ± 3.0 ^a	278.6 ± 15.0 ^a
Cyt-2	Fresh weight	239.6 ± 4.1 ^a	270.0 ± 6.6 ^{ab}	1421.6 ± 76.0 ^a
	Dry weight	59.3 ± 1.5 ^a	87.6 ± 3.7 ^a	277.6 ± 11.2 ^a
Multiple significance is marked with letters by LSD test at $P < 0.05$. Data as means ± SD (n = 3).				

Enlarged stem and leaf organs but reduced pollen activity in polyploid potato plants

Concerning the biomass yields increased in the Tet and Cyt plants, we observed stem and leaf dissections. Compared to the Dip plants, both Tet and Cyt lines exhibited a significant enlargement in different types of stem cells including the sizes of the epidermis, cortex, and xylem cells (Fig. 2a). Furthermore, the Cyt lines showed either significantly increased stem radius and cortex sizes or reduced xylem sizes than those of the Tet plants, but Tet and Cyt plants had a similar epidermis dimension (Table 2). Meanwhile, the Tet and Cyt plants also exhibited distinctively enlarged leaf thickness than the Dip plants, in particular on the spongy tissues (Fig. 2b, Table 2). The data thus revealed that the Cyt plants were of the most enhanced stem and leaf growth, which should be accountable for its highest biomass yield examined above. However, this study found significantly reduced pollen activity in the Cyt and Tet plants, compared to the Dip plants (Table S4).

Table 2
Stem and leaf dissections in the Dip, Tet-1, and Cyt-1 potato samples at the flowering stage.

Samples	Stem (μm)				Leaf (μm)	
	Radius	Epidermis	Cortex	Xylem	Thickness	Spongy tissue
Dip	522.0 \pm 33.0 ^c	19.2 \pm 1.6 ^b	160.0 \pm 2.7 ^c	150.0 \pm 3.5 ^c	96.2 \pm 0.6 ^c	56.3 \pm 0.8 ^c
Tet-1	922.0 \pm 30.0 ^b	31.9 \pm 1.6 ^a	319.0 \pm 7.0 ^b	210.0 \pm 0.6 ^a	102.6 \pm 1.4 ^b	77.6 \pm 0.5 ^b
Cyt-1	1235.0 \pm 30.0 ^a	34.8 \pm 1.6 ^a	541.0 \pm 8.0 ^a	180.0 \pm 2.6 ^b	119.7 \pm 0.8 ^a	81.0 \pm 0.4 ^a

Multiple significance is marked with letters by LSD test at $P < 0.05$. Data as means \pm SD (n = 3).

Relatively stable genotype and phenotype of Cyt by asexual reproduction

Since the cytochimera plant containing two different ploidy cells, to understand whether the cytochimera potato is stable, the genotype status and phenotype stability of Cyt plants were analyzed by using materials from different generations and two different asexual propagations. Percentages of tetraploid cells in the different organs including root, stem, leaf, and flower of Cyt plants at the flowering stage were calculated using 5S rDNA FISH. We found that there is no significant difference in the percentages of tetraploid cells among the four organs at $\alpha = 0.05$ (Table S1). Meanwhile, this study undertook five years to consistently generate more than 30 generations of nodal culture propagations in vitro using the nodal stem segments from original Cyt plants and examined that two Cyt lines of the 30th generation could maintain 30.2 %-33.5 % tetraploid cells (Table S5). Therefore, there was no significant difference in the proportion of tetraploid and diploid cells among generations by χ^2 test, suggesting that the stage of incomplete duplication of Dip genomes could steadily maintain in the Cyt potato plant (Table S5). In addition, Cyt plants propagated by tubers in the field from three continuous generations also showed that the frequencies of tetraploid cells were relatively stable, which kept around one-third of the total cells (Table S1).

The phenotypes of Cyt plants from different generations and two different asexual propagations were also surveyed. We found that the phenotype including plant growth and biomass yields of the two elite lines of each polyploid type (Cyt-1, 2) was also very stable. For example, in the two-years field experiments, the two Cyt lines exhibited the most enhanced plant growth and biomass yields than those of the Dip and Tet plants (Fig. 1b, Table 1).

Altered Cell Wall Composition In Polyploid Potato Straws

Because plant cell walls are the major components of straw biomass, this study determined wall polymers levels in the polyploid potato straws (Table 3). We calculated wall polymers levels against dry matter (% DM), and both cellulose and lignin were significantly reduced in the Tet and Cyt plants relative to the Dip plants, whereas the hemicellulose and pectin were increased (Table 3). In addition, the Cyt

plants showed significantly different contents of four wall polymers from the Tet plants, suggesting that Tet and Cyt plants should have distinct cell wall composition. As pectin and lignin are respectively rich at primary and secondary cell walls (Guo et al. 2014; Hu et al. 2018a; Hu et al. 2018b), the data indicated much raised primary cell wall deposition in the Cyt straw, consistent with its most enhanced plant growth observed in this study.

Table 3

Plant cell wall compositions (% dry matter) of mature straw in the diploid (Dip), tetraploid (Tet), and cytochimera (Cyt) potato samples.

Samples	Cellulose	Hemicellulose	Lignin	Pectin
Dip	19.02 ± 0.54 ^a	9.77 ± 0.21 ^c	19.09 ± 0.87 ^a	6.5 ± 0.13 ^c
Tet-1	16.34 ± 0.42 ^b	11.85 ± 0.18 ^b	15.61 ± 0.65 ^b	7.66 ± 0.21 ^b
Tet-2	16.01 ± 0.20 ^b	11.52 ± 0.43 ^b	16.02 ± 0.30 ^b	7.59 ± 0.25 ^b
Cyt-1	12.20 ± 0.40 ^c	14.30 ± 0.30 ^a	12.90 ± 0.32 ^c	8.87 ± 0.43 ^a
Cyt-2	11.80 ± 0.63 ^c	13.99 ± 0.45 ^a	12.65 ± 0.49 ^c	8.82 ± 0.40 ^a

Multiple significance is marked with letters by LSD test at $P < 0.05$. Data as means ± SD (n = 3).

Characteristic Co-expression Networks Among Ploidy Potato Plants

To analyze why Cyt plants have the highest biomass and significantly altered cell wall polymers, RNA sequencing was performed on three lines with different ploidy. In general, the obtained sequencing data were of good quality at 150 G in total, and the mapping rates with the reference genome were varied around 80 % (Table S6). From the 18 samples examined, the transcripts with a mean FPKM < 2 were filtered and a total of 14584 transcripts were thus retained for the following co-expression network analysis. Because the samples exhibited a similar curve of FPKM distribution (Fig. 3a), we justified that they should be of perfect homogeneity. Meanwhile, this study generated the cluster tree with two classic branches corresponding for stem and leaf tissues, but the cluster was unavailable for the sample types, which indicated that the tissues should be of more specificity for the cluster tree, compared to the sample heterogeneity (Fig. 3b). Furthermore, the Dip samples were of independent branches both in the stem and leaf tissues, whereas Tet and Cyt samples did not show obvious clustering, suggesting that the genome doubling should cause a similar expression pattern between the Tet and Cyt samples. Based on the means of traits (Table S4) described above and in previous research (Madadi et al. 2021), this study generated a heatmap (Fig. 3b) and the Cyt samples were dominant against other samples in almost all traits examined, consistent with the enhanced plant growth and biomass yield in Cyt plants. Then, the co-expression network was constructed to obtain nine distinct functional modules, except the grey module with unclassified transcripts (Fig. 3c).

By correlating the above nine modules with the traits, it was found that the green module was highly correlated with major agricultural traits about overall plant growth and development (Fig. 4a). Using the green module as an example (Fig. 4b), we examined that transcripts of the module were of the high significance measures with the traits and the high module membership with the module eigengenes (MEs). In addition, the line diagrams of MEs showed an expression trend of overall transcripts for each module (Fig. 4c). Despite stem and leaf tissues were of different expression as the fluctuation range of lines, the transcripts within the green module did not show tissue-specific expression, and thus they may mainly respond for polyploidization.

Furthermore, GO and KEGG enrichment analyses were conducted for transcripts of the nine modules (Table 4). The turquoise module exhibited a reducing expression trend in the Tet and Cyt plants (Fig. 4c), which may mainly affect plant cell wall formation and assembly, as well as xylan biosynthesis, consistent with the altered cell wall composition (Table 3). In particular, the green module may associate with spliceosome by participating in pre-mRNA processing. More importantly, this module was highly correlated with many agricultural traits, suggesting that genome doubling can downregulate the expression of spliceosomal associated transcripts and affect transcriptional regulation by changing alternative splicing, thus specifically improving some traits in Tet and Cyt plants (Fig. 4, Table 4). In summary, although the expression trends of Cyt and Tet were similar, such as the turquoise module containing the most transcripts and the green module highly correlated with traits. However, the differential expression of transcripts for lipid transport (pink module), fatty acid biosynthesis (brown module), and steroid biosynthesis (black module) may be related to the higher biomass of Cyt than that of Tet (Fig. 4, Table 4).

Table 4
Analysis of GO and KEGG enrichment.

Modules	ID	Description	P-value	FDR
pink	GO:0006869	lipid transport	0.000195	0.008606
	sot00052	Galactose metabolism	0.002788	0.013942
	sot00040	Pentose and glucuronate interconversions	0.014177	0.035442
magenta	GO:0006098	pentose-phosphate shunt	0.00111	0.042948
	sot00030	Pentose phosphate pathway	0.003803	0.04306
	sot00710	Carbon fixation in photosynthetic organisms	0.006625	0.04306
turquoise	GO:0042254	ribosome biogenesis	1.69E-06	0.000328
	GO:0034660	ncRNA metabolic process	1.79E-06	0.000328
	GO:0015031	protein transport	2.41E-06	0.000328
	GO:0006412	translation	4.00E-06	0.000398
	GO:0009832	plant-type cell wall biogenesis	1.64E-05	0.000921
	GO:0045492	xylan biosynthetic process	3.19E-04	0.006981
	GO:0009664	plant-type cell wall organization	3.69E-04	0.007507
	GO:0048193	Golgi vesicle transport	0.002152	0.035748
	sot03010	Ribosome	4.88E-10	5.61E-08
	sot03050	Proteasome	4.40E-05	0.002532
yellow	GO:0009742	brassinosteroid mediated signaling pathway	0.000262	0.020615
	GO:0010605	negative regulation of macromolecule metabolic process	0.000342	0.024274
red	GO:0006260	DNA replication	6.32E-10	7.11E-08
	GO:0006412	translation	6.86E-10	7.11E-08
	GO:0000819	sister chromatid segregation	0.000287	0.016242
	GO:0033044	regulation of chromosome organization	0.000287	0.016242
	sot03010	Ribosome	8.79E-13	5.62E-11
	sot03030	DNA replication	9.15E-07	2.93E-05
	sot00190	Oxidative phosphorylation	6.04E-06	0.000129
blue	GO:0019684	photosynthesis, light reaction	1.17E-08	3.21E-06

Modules	ID	Description	P-value	FDR
	GO:0009658	chloroplast organization	2.16E-07	4.45E-05
	GO:0006662	glycerol ether metabolic process	3.87E-05	0.003981
	GO:0051186	cofactor metabolic process	0.000107	0.009727
	GO:0046148	pigment biosynthetic process	0.000132	0.010858
	sot00195	Photosynthesis	2.16E-15	2.23E-13
	sot00710	Carbon fixation in photosynthetic organisms	1.07E-05	0.000372
	sot00630	Glyoxylate and dicarboxylate metabolism	1.29E-05	0.000372
	sot03010	Ribosome	1.45E-05	0.000372
	sot01200	Carbon metabolism	8.85E-05	0.001822
	sot00860	Porphyrin and chlorophyll metabolism	0.000108	0.001847
	sot00906	Carotenoid biosynthesis	0.002171	0.031951
brown	GO:0006633	fatty acid biosynthetic process	1.20E-10	7.92E-08
	sot00061	Fatty acid biosynthesis	7.74E-05	0.007435
	sot01212	Fatty acid metabolism	0.000715	0.02746
	sot00062	Fatty acid elongation	0.000858	0.02746
black	sot00100	Steroid biosynthesis	1.02E-08	3.96E-07
green	sot03040	Spliceosome	0.000142	0.01195

Finally, this study identified a total of 267 transcripts with an average degree of 65 in the co-expression network of the green module (Fig. 4d). Among them, the *PGSC0003DMT400023450* (Splicing factor 3a) belongs to the spliceosomal pathway sot03040 (Table 4) and the *PGSC0003DMT400023679* is an RNA binding protein with multiple splicing, confirming that the green module is highly associated with the spliceosome. On the other hand, the related transcripts have been examined in other plant species. For instance, *Arabidopsis* homologous gene *AT5G06160* (ATO) of *PGSC0003DMT400023450* has been examined to implicate in pre-spliceosome formation as a novel regulator of gametic cell fate (Moll et al. 2008). In addition, this study also identified a total of 131 transcripts with an average degree of four in the co-expression network of the turquoise module (Fig. 4e). Using the *Arabidopsis* homologous gene *AT2G38080* (LAC4) of *PGSC0003DMT400028629* (Laccase-4) as an example, which plays a major role in catalyzing lignin monomer polymerization (Berthet et al. 2011; Zhao et al. 2013), confirming that the turquoise module is highly associated with the cell wall.

Discussion

Polyploidization has been widely implemented in plant breeding by doubling genomes for improving grain yield or quality and stress resistance in food crops (Bamakhramah et al. 1984). It is also applied to increase biomass yield and to alter cell wall composition in bioenergy crops (Chae et al. 2013). Recently, it has been examined that the tetraploid *Arabidopsis* plants are of almost double dry biomass yields than those of the diploid ones, but the hexaploid and octoploid plants show significantly reduced biomass yields, compared to the tetraploid plants (Corneillie et al. 2019). As an exception, some polyploid plants could not show any higher grain and biomass yields, as compared to their diploid parents. Hence, the characteristic of polyploidization depends on genomic structure, reproduction patterns, ancestral ploidy level, and the purpose for which the crop has been cultivated (Manzoor et al. 2019). Notably, this study has demonstrated a novel strategy to maximize biomass yield in the potato crop by performing incomplete genome doubling. Despite it remains to examine if this strategy is universally applicable for all food and bioenergy crops, the RNA-seq analyses performed in this study have revealed that incomplete genome doubling could provide multiple options to select the crops that are of specifically improved agricultural traits.

Based on the co-expression network analysis, this study has examined that some important transcripts in genome doubled potatoes (Tet, Cyt) are differently expressed compared to the diploid ones (Dip), which may play a dominating role in enhanced biomass yield. Among the nine modules, the green module associated with the plant growth and biomass yield shows a similar expression pattern between the Tet and Cyt plants (Fig. 4c). Alternative splicing is an important mechanism to regulate gene expression and produce proteome diversity, which plays an important role in plant growth and development (Huang et al. 2020). Because the green module was highly associated with spliceosome, this study suggests that alternative splicing may play an important role in the novel traits observed in the Tet and Cyt plants, consistent with the previous report (Fasano et al. 2016). However, data of major phenotypes showed that Cyt was significantly superior to Tet plants, possibly due to different expression of transcripts within modules of lipid transport, fatty acid biosynthesis, and steroid biosynthesis (Fig. 4c, Table 4).

On the other hand, despite the tetraploid cell with a low level (about 30 %) in Cyt plants, we speculate that those cells should be enough to trigger a similar gene expression pattern to the Tet plants. A possible explanation is that sRNA may transfer between different cell layers of Cyt to change the expression of their target genes (Li et al. 2013). Intercellular morphogenetic interactions and cell displacement might have also occurred in the Cyt plants, support the hypothesis of epigenetic and molecular influences in genetically different cells (Hashimoto-Freitas and Nassar 2013). Interestingly, another Cyt line with 52 % tetraploid cells shows the fresh biomass up to 4.1-fold against the Dip (Cyt-1, 2 up to 9.2-fold), probably due to this Cyt with more tetraploid cells. Therefore, the proportion of different ploidy cells may seriously affect the growth and development of Cyt plants, and the incomplete genome doubling could provide a powerful methodology to improve the agricultural traits of interest in plant breeding.

In polyploid crops, genome doubling could cause loss of self-incompatibility and gain of asexual reproduction (Maciver 2019; Robertson et al. 2011). In this study, we examined significantly reduced pollen activity in the polyploid potato plants, and the Cyt and Tet plants showed similar pollen activity (Fig. 3b, Table S4). Nevertheless, this study has avoided such a disadvantage of polyploidization by the propagation of potato tubers, and more than 30 generations of propagation in vitro using nodal stem segments were examined with genetic stability in the Cyt plants. Hence, the incomplete genome doubling should be applicable in perennial bioenergy crops such as *Miscanthus* and switchgrass (Madadi et al. 2021). In addition, this study did not find out significantly raised tuber yields in the Cyt plants, compared to the Dip plants (Fig. 3b, Table S4), which should be consistent with the assumption that incomplete genome doubling could not improve all agricultural traits in the same line. In conclusion, the Cyt with incomplete genome doubling is a promising new strategy to increase biomass for food and bioenergy crops.

Declarations

Data availability

The RNA sequence is available from the National Centre for Biotechnology Information (NCBI) as BioProject ID PRJNA591474.

Acknowledgments

This work was in part supported by the projects of National Natural Science Foundation of China (32070556, 32060503, 31571721, 31670296), the National 111 Project of Ministry of Education of China (B08032), the Inner Mongolia Science and Technology Project (2020GG0080) and Inner Mongolia Natural Science Foundation (2020ZD09).

Authors' contributions

KZ completed the major experiments of potato polyploidy identification and bioinformatic analysis, and wrote the manuscript; NJ, LW, Zhijun X, JW, and JD participated in the potato polyploidy selection and characterization; MM completed the most potato biomass process experiments; Youmei W, ST, and Yanting W participated in the tissue culture, biomass chemical analysis, and experiment discussion; LP supervised biomass process project and finalized the manuscript, and Zhiyong X designed and supervised the potato polyploidy project.

Conflicts of interest

The authors declare no conflict of interest.

Ethics approval

Not applicable.

Consent to participate

Not applicable.

Consent for publication

Not applicable

References

1. Adams KL, Wendel JF (2005) Novel patterns of gene expression in polyploid plants. *Trends Genet* 21:539-543
2. Auger DL, Gray AD, Ream TS, Kato A, Coe EH, Birchler JA (2005) Nonadditive gene expression in diploid and triploid hybrids of maize. *Genetics* 169:389-397
3. Bamakhramah H, Halloran G, Wilson J (1984) Components of yield in diploid, tetraploid and hexaploid wheats (*Triticum* spp.). *Ann Bot*:51-60
4. Berthet S, Demont-Caulet N, Pollet B, Bidzinski P, Cezard L, Le Bris P, Borrega N, Herve J, Blondet E, Balzergue S, Lapierre C, Jouanin L (2011) Disruption of LACCASE4 and 17 results in tissue-specific alterations to lignification of *Arabidopsis thaliana* stems. *Plant Cell* 23:1124-1137
5. Cai D, Chen J, Chen D, Dai B, Zhang W, Song Z, Yang Z, Du C, Tang Z, He Y, Zhang D, He G, Zhu Y (2007) The breeding of two polyploid rice lines with the characteristic of polyploid meiosis stability. *Sci China C Life Sci* 50:356-366
6. Carputo D, Frusciante L, Peloquin SJ (2003) The role of 2n gametes and endosperm balance number in the origin and evolution of polyploids in the tuber-bearing solanums. *Genetics* 163:287-294
7. Chae WB, Hong SJ, Gifford JM, Rayburn AL, Widholm JM, Juvik JA (2013) Synthetic polyploid production of *Miscanthus sacchariflorus*, *Miscanthus sinensis*, and *Miscanthus x giganteus*. *Global Change Biology Bioenergy* 5:338-350
8. Comai L (2005) The advantages and disadvantages of being polyploid. *Nat Rev Genet* 6:836-846
9. Corneillie S, De Storme N, Van Acker R, Fangel JU, De Bruyne M, De Rycke R, Geelen D, Willats WGT, Vanholme B, Boerjan W (2019) Polyploidy Affects Plant Growth and Alters Cell Wall Composition. *Plant Physiol* 179:74-87
10. Dische Z, Pallavicini C, Kawasaki H, Smirnow N, Cizek LJ, Chien S (1962) Influence of the nature of the secretory stimulus on the composition of the carbohydrate moiety of glycoproteins of the submaxillary saliva. *Arch Biochem Biophys* 97:459-469
11. Dong F, Song J, Naess S, Helgeson J, Gebhardt C, Jiang J (2000) Development and applications of a set of chromosome-specific cytogenetic DNA markers in potato. *Theor Appl Genet* 101:1001-1007
12. Fasano C, Diretto G, Aversano R, D'Agostino N, Di Matteo A, Frusciante L, Giuliano G, Carputo D (2016) Transcriptome and metabolome of synthetic *Solanum* autotetraploids reveal key genomic stress events following polyploidization. *New Phytol* 210:1382-1394

13. Frank MH, Chitwood DH (2016) Plant chimeras: The good, the bad, and the 'Bizzaria'. *Dev Biol* 419:41-53
14. Fry SC, Fry S, Fry S, Fry SC, Fry C (1988) *The Growing Plant Cell Wall: Chemical and Metabolic Analysis*. Growing Plant Cell Wall Chemical & Metabolic Analysis
15. Fujishige I, Tanaka R, Taniguchi K (1996) Efficient isolation of non-chimeric tetraploids artificially induced in a stable culture of *Haplopappus gracilis*. *Theor Appl Genet* 92:157-162
16. Guo K, Zou WH, Feng YQ, Zhang ML, Zhang J, Tu F, Xie GS, Wang LQ, Wang YT, Klie S, Persson S, Peng LC (2014) An integrated genomic and metabolomic framework for cell wall biology in rice. *BMC Genomics* 15
17. Ha M, Lu J, Tian L, Ramachandran V, Kasschau KD, Chapman EJ, Carrington JC, Chen XM, Wang XJ, Chen ZJ (2009) Small RNAs serve as a genetic buffer against genomic shock in *Arabidopsis* interspecific hybrids and allopolyploids. *PNAS* 106:17835-17840
18. Hashimoto-Freitas DY, Nassar NM (2013) Cytogenetic and anatomic behavior of cytochimeras and total polyploids in cassava. *Genet Mol Res* 12:4879-4894
19. Hu HZ, Zhang R, Feng SQ, Wang YM, Wang YT, Fan CF, Li Y, Liu ZY, Schneider R, Xia T, Ding SY, Persson S, Peng LC (2018a) Three *AtCesA6*-like members enhance biomass production by distinctively promoting cell growth in *Arabidopsis*. *Plant Biotechnol J* 16:976-988
20. Hu HZ, Zhang R, Tao ZS, Li XK, Li YY, Huang JF, Li XX, Han X, Feng SQ, Zhang GM, Pen LC (2018b) Cellulose Synthase Mutants Distinctively Affect Cell Growth and Cell Wall Integrity for Plant Biomass Production in *Arabidopsis*. *Plant and Cell Physiology* 59:1144-1157
21. Huang J, Lu XY, Wu HW, Xie YC, Peng Q, Gu LF, Wu JY, Wang YC, Reddy ASN, Dong SM (2020) *Phytophthora* Effectors Modulate Genome-wide Alternative Splicing of Host mRNAs to Reprogram Plant Immunity. *Mol Plant* 13:1470-1484
22. Langfelder P, Horvath S (2008) WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 9
23. Leitch AR, Leitch IJ (2008) Perspective - Genomic plasticity and the diversity of polyploid plants. *Science* 320:481-483
24. Li JX, Wang Y, Zhang LL, Liu B, Cao LW, Qi ZY, Chen LP (2013) Heritable variation and small RNAs in the progeny of chimeras of *Brassica juncea* and *Brassica oleracea*. *J Exp Bot* 64:4851-4862
25. Maciver SK (2019) Ancestral Eukaryotes Reproduced Asexually, Facilitated by Polyploidy: A Hypothesis. *Bioessays* 41:e1900152
26. Madadi M, Zhao K, Wang Y, Wang Y, Tang S-w, Xia T, Jin N, Xu Z, Li G, Qi Z, Peng L, Xiong Z (2021) Modified lignocellulose and rich starch for complete saccharification to maximize bioethanol in distinct polyploidy potato straw. *Carbohydr Polym* 265:118070
27. Manzoor A, Ahmad T, Bashir MA, Hafiz IA, Silvestri C (2019) Studies on Colchicine Induced Chromosome Doubling for Enhancement of Quality Traits in Ornamental Plants. *Plants (Basel)* 8

28. Moll C, von Lyncker L, Zimmermann S, Kagi C, Baumann N, Twell D, Grossniklaus U, Gross-Hardt R (2008) CLO/GFA1 and ATO are novel regulators of gametic cell fate in plants. *Plant J* 56:913-921
29. Ortiz R, Ehlenfeldt MK (1992) The importance of Endosperm Balance Number in potato breeding and the evolution of tuber-bearing *Solanum* species. *Euphytica* 60:105-113
30. Otto SP (2007) The evolutionary consequences of polyploidy. *Cell* 131:452-462
31. Parisod C, Holderegger R, Brochmann C (2010) Evolutionary consequences of autopolyploidy. *New Phytol* 186:5-17
32. Peng L, Hocart CH, Redmond JW, Williamson RE (2000) Fractionation of carbohydrates in *Arabidopsis* root cell walls shows that three radial swelling loci are specifically involved in cellulose production. *Planta* 211:406-414
33. Pertea M, Kim D, Pertea GM, Leek JT, Salzberg SL (2016) Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat Protoc* 11:1650-1667
34. Robertson K, Goldberg EE, Igic B (2011) Comparative evidence for the correlated evolution of polyploidy and self-compatibility in Solanaceae. *Evolution* 65:139-155
35. Sattler MC, Carvalho CR, Clarindo WR (2016) The polyploidy and its key role in plant breeding. *Planta* 243:281-296
36. Serapiglia MJ, Gouker FE, Hart JF, Unda F, Mansfield SD, Stipanovic AJ, Smart LB (2015) Ploidy Level Affects Important Biomass Traits of Novel Shrub Willow (*Salix*) Hybrids. *Bioenergy Res* 8:259-269
37. Shelley, Jansky, David, Spooner (2018) The Evolution of Potato Breeding. *Plant Breeding Reviews*
38. Sluiter A, Hames B, Ruiz R, Scarlata C, Sluiter J, Templeton D, Crocker D (2008) Determination of structural carbohydrates and lignin in biomass. *Laboratory analytical procedure* 1617:1-16
39. Song K, Lu P, Tang K, Osborn TC (1995) Rapid genome change in synthetic polyploids of *Brassica* and its implications for polyploid evolution. *PNAS* 92:7719-7723
40. Stupar RM, Bhaskar PB, Yandell BS, Rensink WA, Hart AL, Ouyang S, Veilleux RE, Busse JS, Erhardt RJ, Buell CR, Jiang JM (2007) Phenotypic and transcriptomic changes associated with potato autopolyploidization. *Genetics* 176:2055-2067
41. Wendel JF (2000) Genome evolution in polyploids. *Plant Mol Biol* 42:225-249
42. Xiong Z, Gaeta RT, Edger PP, Cao Y, Zhao K, Zhang S, Pires JC (2021) Chromosome inheritance and meiotic stability in allopolyploid *Brassica napus*. *G3* 11:jkaa011
43. Xiong ZY, Gaeta RT, Pires JC (2011) Homoeologous shuffling and chromosome compensation maintain genome balance in resynthesized allopolyploid *Brassica napus*. *PNAS* 108:7908-7913
44. Yu GC, Wang LG, Han YY, He QY (2012) clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. *Omics-a Journal of Integrative Biology* 16:284-287
45. Zhang J, Liu Y, Xia EH, Yao QY, Gao LZ (2015) Autotetraploid rice methylome analysis reveals methylation variation of transposable elements and their effects on gene expression. *Proc Natl Acad Sci U S A* 112:201515170

46. Zhao Q, Nakashima J, Chen F, Yin YB, Fu CX, Yun JF, Shao H, Wang XQ, Wang ZY, Dixon RA (2013) LACCASE Is Necessary and Nonredundant with PEROXIDASE for Lignin Polymerization during Vascular Development in Arabidopsis. *Plant Cell* 25:3976-3987

Figures

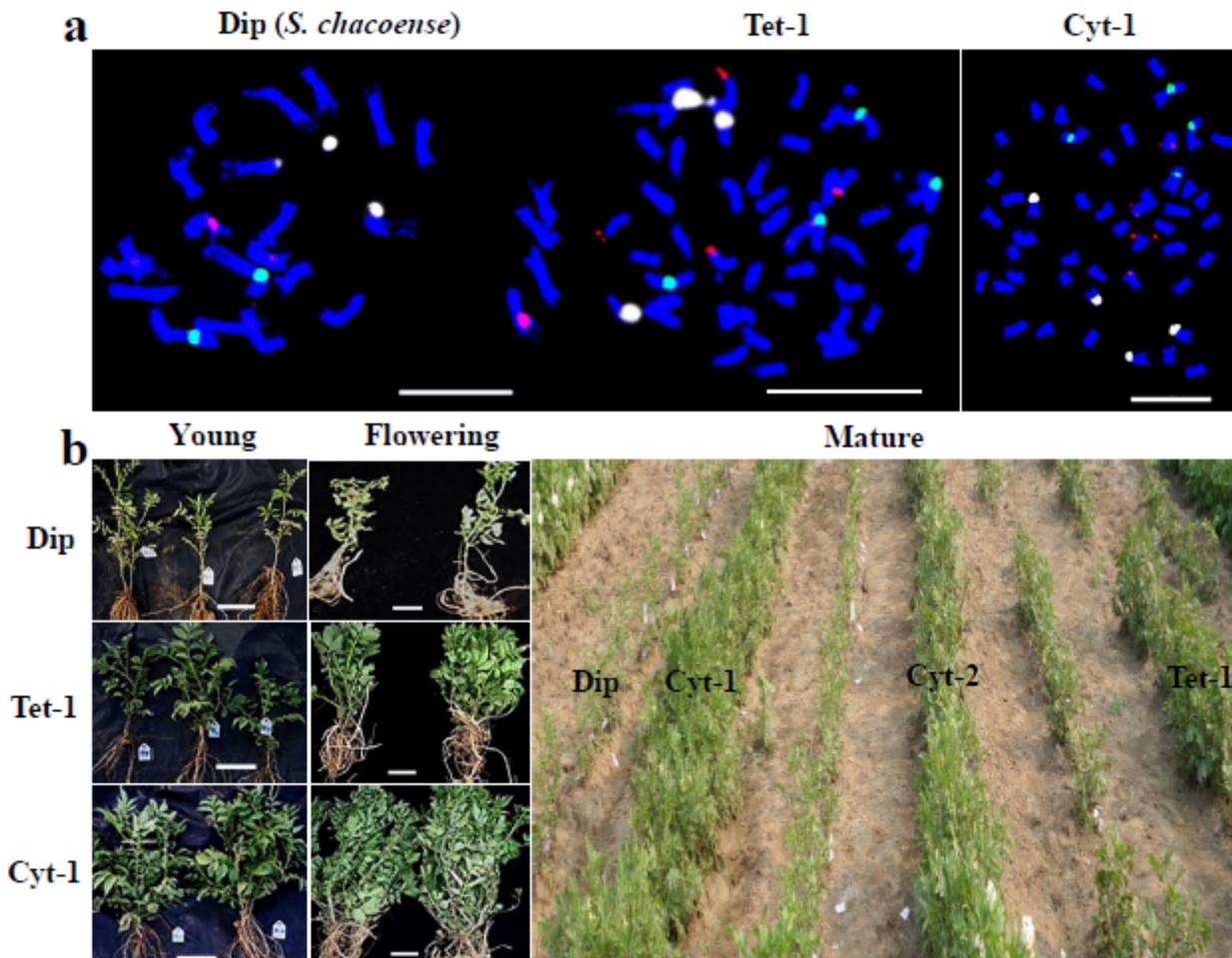


Figure 1

Ploidy identification and growth of potatoes. (a) Identification of the chromosome integrality of diploid (Dip), tetraploid (Tet), and cytochimera (Cyt) samples by FISH with 5S rDNA (green, located in chromosome 1), 45S rDNA (white, located chromosome 2) and chromosome 3 specific BAC 079E02 (red) as probes. Dip potato gets 24 chromosomes including one pair of chromosome 1, 2 and 3, while Tet-1 contains 48 chromosomes including two pairs of chromosome 1, 2 and 3. Cyt-1 has both diploid and tetraploid cells, and the tetraploid cell of Cyt-1 has 48 chromosomes consisting of two pairs of chromosome 1, 2 and 3. Scale bar, 10 μ m. (b) Plant growth of Dip, Tet-1, and Cyt-1 at three different stages in field. Scale bar, 20 cm.

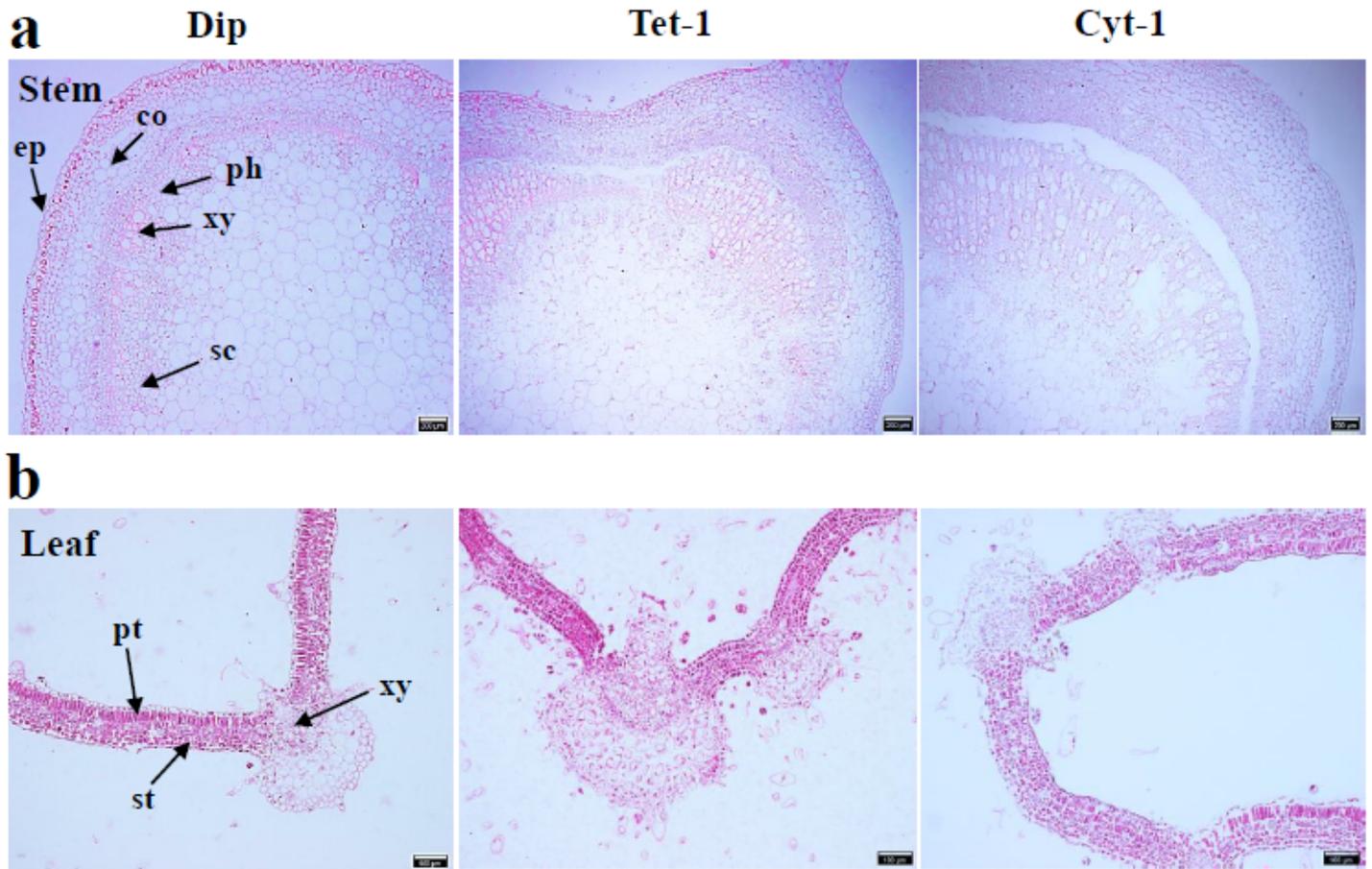


Figure 2

Cross-section observation of stem and leaf in the Dip, Tet-1, and Cyt-1 potato samples at the flowering stage. Images of the stem (a) and leaf (b) dissection under microscopy. ep, epidermis; co, cortex; ph, phloem; xy, xylem; sc, sclerenchyma cell; pt, palisade tissue; st, spongy tissue. Scale bar, 200 μm (stem) and 100 μm (leaf).

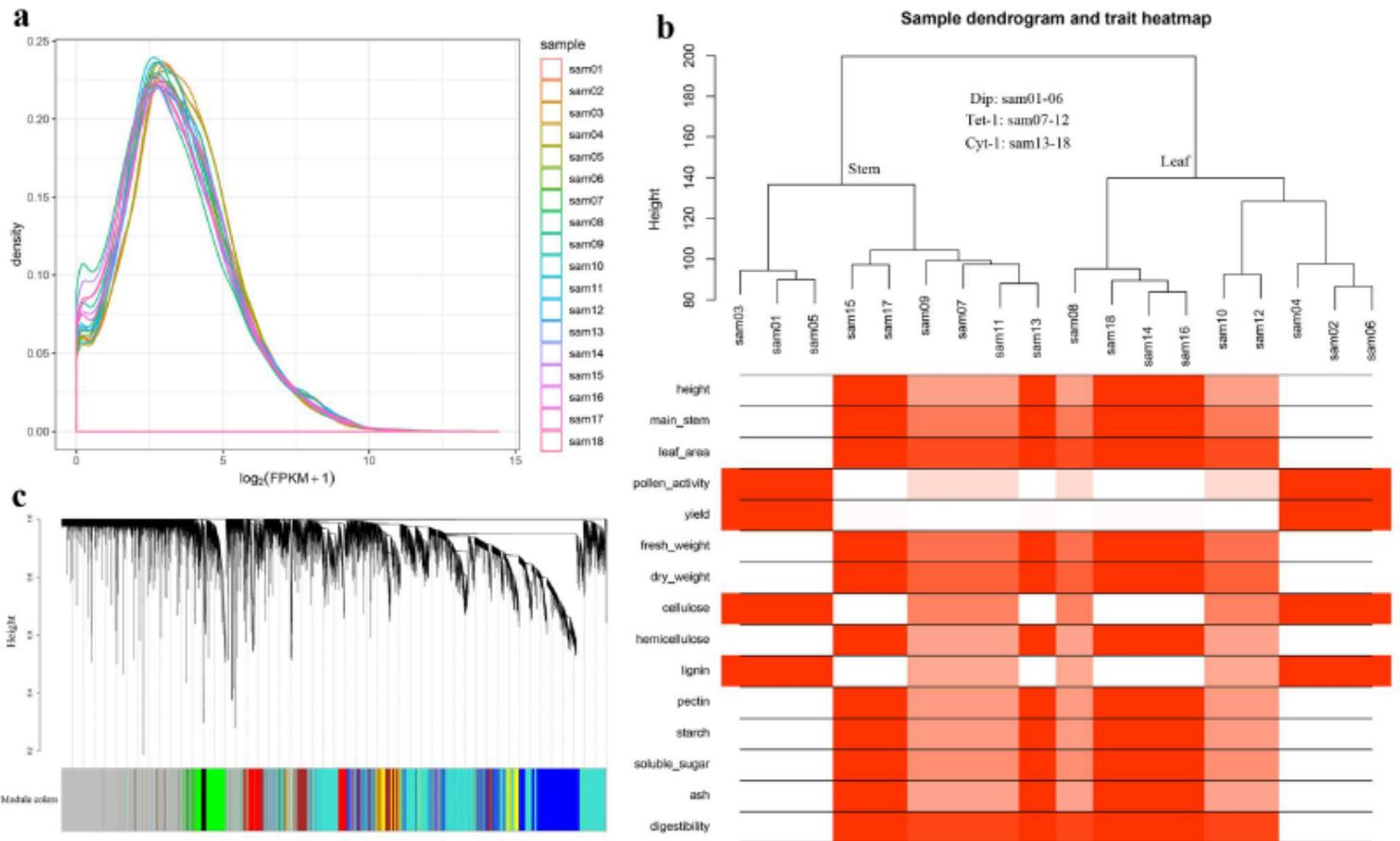


Figure 3

Potato samples clustering and traits heatmap. (a) The original FPKM is standardized by $\log_2(\text{FPKM}+1)$ for the density distribution curve. (b) In the clustering tree, the odd and even numbers represent the samples from stem and leaf, relatively. Sam01-06, sam07-12, and sam13-18 are samples of Dip, Tet-1, and Cyt-1, respectively. In the heatmap of potato traits, the trait values from high to low corresponded to the heatmap's color from red to white. (c) Transcripts clustering to divide modules.

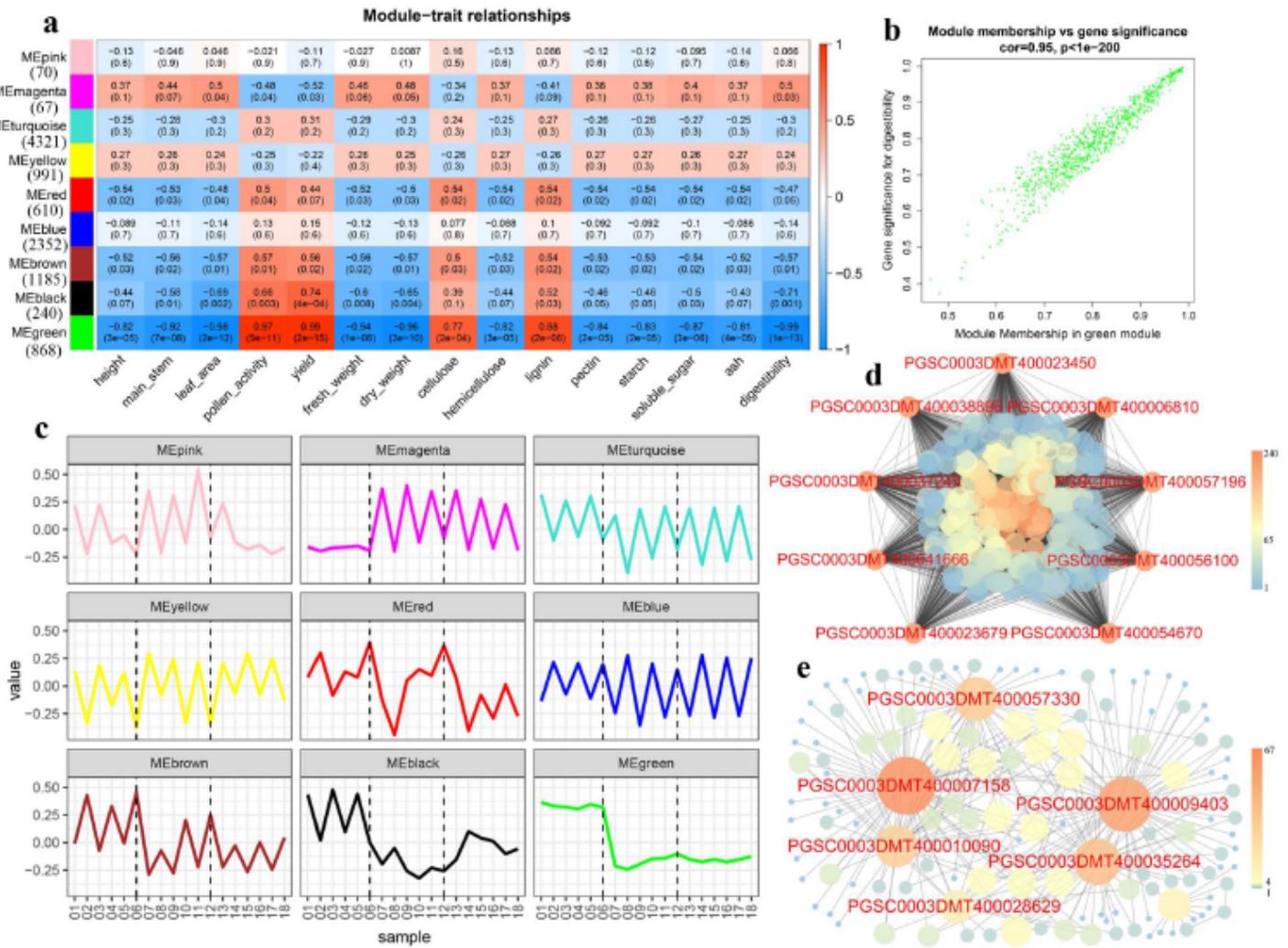


Figure 4

WGCNA of the Dip, Tet-1, and Cyt-1 samples. (a) The relationship of modules and traits. The numbers below the modules' name are the number of transcripts contained, and each square shows the correlation (P-value). The red and blue backgrounds are of positive and negative correlations, respectively. (b) Scatterplots of gene significance versus module membership for the digestibility with MEgreen. Significant correlations imply that hub transcripts within the modules are also very correlated with the traits. (c) Expression trend of overall transcripts within the modules. The odd and even numbers represent the samples from stem and leaf, relatively. Sam01-06, sam07-12, and sam13-18 are samples of Dip, Tet-1, and Cyt-1, respectively. (d) Co-expression network of the green module. In the network, edge weight ≥ 0.15 , and the higher the connection degree is, the deeper the color. The top nine dots (hub transcripts) show the name in red. (e) Co-expression network of the turquoise module. In the network, edge weight ≥ 0.30 , and the higher the connection degree is, the deeper the color and the bigger the dot. The top six dots (hub transcripts) show the name in red.

Supplementary Files

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

- [SupplementaryInformation.pdf](#)