

Balanced duplicated gene expression supports an autotetraploid ancestor of Salicaceae plants

Yu-Hao Zhao

North China University of Science and Technology

Jin-Peng Wang

North China University of Science and Technology

Jia-Qing Yuan

North China University of Science and Technology

Jing Li

North China University of Science and Technology

Wei-Na Ge

North China University of Science and Technology

Jin Zhang

North China University of Science and Technology

Yangqin Xie

North China University of Science and Technology

Yue Hou

North China University of Science and Technology

Tariq Shehzad

University of West Georgia

Xiyin Wang (✉ wang.xiyin@gmail.com)

North China University of Science and Technology

Research article

Keywords: Poplar, Autopolyploid, Gene colinearity, Gene expression, Gene family

Posted Date: April 17th, 2019

DOI: <https://doi.org/10.21203/rs.2.9236/v1>

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

Abstract

Background Autopolyploids refer to the increase in the genome from the same species, usually produced by direct doubling of diploid chromosomes. The polyploid formed by chromosome doubling of the same species is called homologous polyploid. Results In order to further check whether the Salicaceae-common tetraploid is homologous or heterologous, with grape as the outer group, by performing gene collinearity analysis, we explored whether two sets of poplar chromosomes or chromosomal regions have balanced gene expression levels and similar gene function. Paired T-test showed that duplicated genes in collinearity were balanced in expression, which is expected if the tetraploid ancestor was homologous whole-genome duplication, or autopolyploidization. Moreover, KEGG enrichment analysis and pathway annotation showed that most of the differentially expressed genes were related to metabolism. A comparison of different groups of flowering plants suggests that autopolyploidization may not provide comparable biological and evolutionary vigor to establish large plant groups, as observed in poaceae and brassicaceae families. The present analysis contributes to understanding the biology and evolution of Salicaceae plants and beyond. Conclusions There was no significant difference in gene expression and gene function between two sets of genomes of poplar.

Background

Poplar (*Populus trichocarpa*), a Salicaceae plant, can be used as fuel and also an important raw material for manufacturing industry [1]. Burning non-renewable fossil fuels has led to massive accumulation of CO₂ in the atmosphere and global warming [2]. Poplar was the first perennial woody plant to have had its whole genome sequence deciphered [3].

Polyploidization plays an important role in the evolution of plants [4, 5]. Previous studies suggested that polyploidy may be an important factor in the formation of large land plant groups [6, 7]. There is clear evidence that one angiosperm genome has undergone at least one whole-genome doubling. For example, the monocot gramineae plants underwent a whole genome-doubling ~100 million years ago (MYA)[5]. Major eudicots are thought to have originated from a common ancient hexaploid plant ~115-130 MYA[7]. Polyploidization doubles or triples chromosomes and genes on them, which in turn undergo complex genetic rearrangements, frequent homoeologous recombination , enormous gene losses and other genetic changes [8]. These resulted in structural and functional complexity of extant plant genomes.

The use of next-generation sequencing technologies by the plant research community has led to the accumulation of large volumes of expression data. However, analysis of these datasets is complicated by the frequent occurrence of polyploidy during the evolution of plants [9]. As a supposed ancient allotetraploid, maize was found to have two diverged subgenomes, with one being dominant and the other sensitive, and the genes from the dominant subgenome have higher expression levels than their duplicated copies from the sensitive one [10]. Moreover, the dominant subgenome preserved more genes than the sensitive one.

With the sequencing of poplar genome, a specific genome doubling event was inferred to have occurred 60 MYA, shared with other salicaceae plants [3]. Having a conservative genome resembling that of the common ancestor of dicot plants, grape is often taken as a reference to understand the structures and evolution of other dicot plant genomes [11]. By referring to gene collinearity, a comparison to grape genome showed that ~68.7% duplicated genes were removed from duplicated regions, deleted or translocated to other genomic regions [12]. Actually, poplar duplicated blocks, especially the larger ones, often have similar gene retention/loss rates [12]. That is, they could not be separated into two different groups or subgenomes, being dominant or sensitive, raising the likelihood of autotetraploidy nature of the genome doubling event [12]. However, the expression profiles of duplicated genes have not been evaluated so far, which would add more knowledge to understanding the nature of the event and the divergence between duplicated genes.

Polyploidization produced thousands of duplicates in extant genomes, even after often wide-spread gene losses. These duplicates would change the biological functions, enhance or reduce gene expression or translation, and rewire pathways and/or networks that they engage. A research into C4 pathway found that duplicated genes produced by ~100-MYA tetraploidization in grass common ancestor constituted a reservoir being a pre-condition for the gradual transit from C3 state to the establishment of C4 function [13]. In addition, the Solanaceae lineage has experienced two consecutive genome triplications: one that is ancient and shared with rosids, and a more recent one. These triplications set the stage for the neofunctionalization of genes controlling fruit characteristics, such as colour and fleshiness [14].

Calcium-dependent protein kinases(CDPKs) are important Ca²⁺sensors and effectors in Ca²⁺ signal transduction pathway. They play crucial roles in regulation of plant development and tolerance of various environmental stresses. Gene expression profiling showed that a number of *Populus* CDPK differentially expressed across different tissues and developmental stages [15].

Here, we checked the gene expression of poplar duplicated genes in the tetraploid event, and whether the event affected the copy number changes of the CDPK gene family, which may help understand evolutionary and functional effects due to ancient polyploidization.

Results

Balanced gene expression of tetraploidy-produced genes

To find whether the tetraploidy-produced genes have similar or divergent expression levels, we inferred 4713 pairs of tetraploidy-produced genes with grape orthologs. The present inferred gene pairs are an update of our previous work performed. These poplar pairs of genes and their grape orthologs are all in collinearity, supporting their origination by the tetraploid event. We distinguished the two sub-genomes of poplar tetraploid events according to grape chromosome numbers and identified them with purple rectangle and green rectangle(Supplementary Figure 1).

We downloaded expression data of five tissues (xylem, phloem, shoot, leaf, and root) and two wood forming cell types (fiber and vessel) of poplar [17]. Each of the tissues and cell types has three repeats. In total, we have 24 datasets. Using paired sample T-test, for each dataset, we compared expression of duplicated copies, which were naturally classified into 19 groups as to from which chromosome are their grape orthologs located (**Table 1**).

In each dataset of studied tissues, for each group, we found that duplicated genes are of the same expression levels. Actually, of the total 456 comparisons, only a tiny fraction of 3.1% (or 14) show significant different expression (P-value < 0.05). This means that the expressional divergence occurs at a mere 3.0% of all comparisons. Besides, if the significance level is set to be 0.01, none comparison shows divergent expression.

Gene expression and gene retention

We checked whether gene expression is related to gene retention. We counted numbers of retained poplar collinear genes as to 19 grape chromosomes, and inferred the average expression of these collinear genes. By performing Pearson correlation test, we found that there was moderate correlation between the retention of collinear genes and the expression level (**Figure 1**; Pearson coefficient = 0.269, P-value = 0.102).

KEGG and GO analysis of duplicates

We performed KEGG of all poplar genes (**Supplementary Table 1**). Of the 4,713 tetraploidy-produced duplicates, actually, only 29 pairs have been annotated with both copies differentially annotated, accounting up to 0.6% of all annotated pairs

Among 4713 pairs of poplar genes, we screened the top 500 pairs of genes with the greatest difference in expression by T-test, and we performed KEGG enrichment analysis over them (**Figure 2**). We found that the most significant enrichment was about carbon fixation in photosynthesis, and then carbon metabolism and fructose and mannose metabolism. All of these are about the most important functions of life support activities.

As to the pathway annotation (**Figure 3**), gene pairs mainly concentrated in metabolism, genetic information processing, environmental information processing, cellular processes, and organismal systems. Among them, the genes concentrated on metabolism were the most, and 403 genes were distributed in 10 sub-functions such as carbohydrate metabolism, energy metabolism, amino acid metabolism, etc. Other functional pathways involve smaller number of genes, for example, genetic information processing: 99 genes, cellular processes: 39 genes, organismal systems: 14, environmental information processing: 19 genes, respectively.

As to GO analysis, duplicated regions did not show any divergence about gene ontology enrichment. For example, we checked poplar duplicated regions orthologous to grape chromosome 2, the collinear genes preserved show no difference in GO item enrichment (**Figure 4**).

Tetraploidization and gene evolution

An example of gene family can help understand gene copy number variation, gene loss and divergent evolutionary rate. Calcium-dependent protein kinases (CDPKs) play crucial roles in regulation of plant development and tolerance of various environmental stresses. Gene expression profiling showed that a number of *Populus* CDPK differentially expressed across different tissues and developmental stages. So we downloaded the CDPK gene family sequence of *Arabidopsis thaliana* and searched and retrieved their homologs in the poplar and grape genomes. These poplar and grape genes were constructed into phylogenetic trees by MEGA(**Figure 5**). There are 9 genes in grape, and 17 genes in poplar, showing a near doubled number of CDPK genes in poplar as to grape. Actually, all these genes are in colinearity within/between genomes, suggesting the copy number increase in poplar is a direct outcome of its specific tetraploidization. No recent tandem duplication was found.

There is clear evidence of genome fractionation by gene loss. At least three poplar paralogs (of pt16G00564, pt05G01136, pt14G01035, respectively) were lost after the tetraploidization, and one grape gene orthologous to pt16G01172 and pt06G01013 was lost. There are six subgroups each with a grape gene and two corresponding poplar orthologs, duplicated in the poplar tetraploidization.

In five out of six subtrees, as expected the grape gene is the outgroup of the poplar duplicates. However, there is one subtree, in which a poplar duplicate is outgroup to the grape gene and the other poplar duplicate, showing an aberrant subtree topology. This can be explained by elevated evolutionary rate in the poplar duplicate coming to be the outgroup.

Disscussion

Previously, the analysis of Salicaceae-common tetraploidization suggested balanced gene losses between the duplicated regions its autoploidization [12]. Here, by analyzing gene expression data, GO and KEGG pathway, we obtained duplicated regions did not show any divergence about gene ontology enrichment, providing further lines of evidence that the event may have an autotetraploidization nature. Our analysis came to results that are distinct from those with maize, which found divergent gene loss rates and expression levels between duplicated regions, suggesting an allotetraploid nature of maize-specific whole-genome duplication [25]. The present findings lay a foundation for further analysis of Salicaceae genomes and understand their biology.

More and more evidence shows that polyploidization has recursively affected the evolution of land plants, and possibly explains their fast divergence and establishment of large plant groups nature [8, 26] . Core eudicots share a hexaploid ancestor living ~130 million years ago [27]. At the family level, the largest plant families of grasses [28], brassicaceae [29], fabaceaeas [30], solanaceaeas [31], and cucurbitaceaeas [32] each share at one family-common polyploidization tens of million years ago , suggesting that polyploidization may be the cause of their formation and success on this plants. Polyploidization can be genetically classified as autoploidization, resulting from the duplication of the same genomes, or allopolyploidization, resulting from the hybridization of diverged genomes [33]. There have been evidences that brassicaceas and poaceaes each have an allopolyploidization ancestor [34, 35], whereas for the other families it is still elusive about the nature of their respective family-common polyploidization. Anyway, allopolyploidization may bring energy to speciation and

divergence due at least partly to hybridization [36]. Recently, it was proposed that soybean [30, 37], kiwifruit [38], salicaceae plants[12] (as further discussed here) may each have an autopolyploid plants tens of millions years ago. Interestingly, we found that though the autopolyploid events are about similarly old, they have produced only small plant groups. This implies that autopolyploidization may not provide comparable opportunities as to allopolloidization to a plant. The mechanisms underlying the observation may be related to hybridization, biological vigor, genome depression, etc, and have to be left for further investigation.

Materials And Methods

Materials

The whole genome sequences of grape and poplar was downloaded from JGI website (<https://genome.jgi.doe.gov/portal/>) [16]. Gene expression data were downloaded from NCBI website (<https://www.ncbi.nlm.nih.gov/>) [17].Arabidopsis CDPK gene is downloaded from TAIR database (www.arabidopsis.org/)

Methods

Inferred gene collinearity.

Gene collinearity is precious to understand ancestral genome structure and evolution. Colinear genes are those genes having preserved ancestral gene orders in extant genomes, which may be often intervened by other genes without collinear counterparts, aroused due to gene loss, relocation, or insertion. Here, to find collinear genes, we inferred putative homologous genes by using BlastP (E_value 1e-5)[18]. A loose definition of gene homology here will not jeopardize the aim of understanding genome structure, but accommodate often much diverged duplicated gene sequences after tens of million years. By using ColinearScan [19] implementing a dynamic programming algorithm (Parameter: non-colinear genes in gap between collinear genes = 40), we inferred colinear gene blocks in poplar and between poplar and grape. By using Nei-Gojobori approach implemented in PAML [20], we estimated synonymous nucleotide substitution rates or Ks between the above inferred collinear genes. Then with the gene collinearity and Ks information, we constructed Dotplot map between poplar and grape, and showed mean Ks between poplar-grape collinear gene blocks.

Gene expression analysis

We downloaded expression data of five tissues (xylem, phloem, shoot, leaf, and root) and two wood forming cell types (fiber and vessel) of poplar [17]. Genes with at least two counts-per-million (cpm) in at least three samples were retained and normalized using the trimmed mean of M value (TMM)[21]. We used grape as a reference. Based on 19 grape chromosomes, we tested whether there were expressional differences between the two sets of poplar genomes by using paired T-test. .

GO and KEGG analysis

As to gene function analysis, we retrieved their GO annotation and KEGG annotation for poplar genes[22, 23]. Because the KEGG annotations are almost all single annotations, we compared the differences of KEGG annotations between the tetraploidy-produced duplicated genes in poplar, and did GO functional enrichment analysis by checking GO functional annotations (default parameters). Then, we screened the top 500 genes with the greatest difference in expression by T-test, and performed KEGG enrichment analysis and KEGG pathway annotation.

Gene family analysis

As to gene family analysis, We downloaded all the genes of the CDPK family in Arabidopsis thaliana (<https://www.arabidopsis.org/>), and then by referring to Arabidopsis CDPK family genes we searched and retrieved poplar CDPK genes through Blast. Then we constructed a phylogenetic tree of poplar CDPK gene family by running MEGA[24].

Declarations

Acknowledgements

We appreciate financial support from the Ministry of Science and Technology of the People's Republic of China (2016YFD0101001), China National Science Foundation (31510333 to J.W. and 3117022 to X.W.), and Natural Science Foundation of Hebei Province (C2015209069 to J.W.) and Hebei New Century 100 Creative Talents Project, Hebei 100 Talented Scholars project, and Tangshan Key Laboratory Project to X.W. We thank the helpful discussion with researchers at the iGeno Co. Ltd., China

Author Contribution

XYW, JPW, and TS conceived the work. YHZ, JQY, JL performed the analysis. JPW, WNG, JZ, YQX.YH performed the analysis with constructive discussions. XYW, JPW, TS, and YHZ prepared the paper.

Declaration of Interest

The authors declare no conflict of interest.

Reference

1. Wang YX, Xu LA, Huang MR, Xu Y: [Progress in research on forest tree comparative genomics]. *Hereditas* 2007, **29**(10):1199.
2. Vitousek PM, D'Antonio CM, Loope LL, Rejmánek M, Westbrooks R: INTRODUCED SPECIES: A SIGNIFICANT COMPONENT OF HUMAN-CAUSED GLOBAL CHANGE. *New Zealand Journal of Ecology* 1997, **21**(1):1-16.
3. Tuskan G A, Difazio S, S J, J B, I G, U H, N P, S R, S R, A S *et al*: The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* 2006, **313**(5793).
4. Paterson AH, Bowers JE, Chapman BA: Ancient polyploidization predating divergence of the cereals, and its consequences for comparative genomics. *Proceedings of the National Academy of Sciences of the United States of America* 2004, **101**(26):9903-9908.
5. Wang, Xiyin, Wang, Jingpeng, Jin, Dianchuan, Guo, Hui, Lee, TaeHo: Genome Alignment Spanning Major Poaceae Lineages Reveals Heterogeneous Evolutionary Rates and Alters Inferred Dates for Key Evolutionary Events. 2015, **8**(6):885-898.
6. DE S, CD B, S K, PS S: Origin and early evolution of angiosperms. *Annals of the New York Academy of Sciences* 2010, **1133**(1):3-25.
7. Jiao Y, Wickett NJ, Ayyampalayam S, AS C, L L, PE R, LP T, Y H, H L, PS S: Ancestral polyploidy in seed plants and angiosperms. *Nature* 2011, **473**(7345):97-100.
8. Soltis PS, Marchant DB, Van dPY, Soltis DE: Polyploidy and genome evolution in plants. *Current Opinion in Genetics & Development* 2015, **35**(2):119-125.
9. Pearce S, Vazquez-Gross H, Herin SY, Hane D, Wang Y, Gu YQ, Dubcovsky J: WheatExp: an RNA-seq expression database for polyploid wheat. *Bmc Plant Biology* 2015, **15**(1):1-8.
10. James C S, Nathan M S, Michael F: Differentiation of the maize subgenomes by genome dominance and both ancient and ongoing gene loss. *Proceedings of the National Academy of Sciences of the United States of America* 2011, **108**(10):4069-4074.
11. Jiao Y, Leebens-Mack J, Ayyampalayam S, Bowers JE, McKain MR, McNeal J, Rolf M, Ruzicka DR, Wafula E, Wickett NJ *et al*: A genome triplication associated with early diversification of the core eudicots. *Genome Biology* 2012, **13**(1):R3.
12. Liu Y, Wang J, Ge W, Wang Z, Li Y, Yang N, Sun S, Zhang L, Wang X: Two Highly Similar Poplar Paleo-subgenomes Suggest an Autotetraploid Ancestor of Salicaceae Plants. *Frontiers in Plant Science* 2017, **8**(R3).
13. Wang X, Gowik U, Tang H, Bowers JE, Westhoff P, Paterson AH: Comparative genomic analysis of C4 photosynthetic pathway evolution in grasses. *Genome Biology* 2009, **10**(6):R68.
14. Consortium TTG: The tomato genome sequence provides insights into fleshy fruit evolution. *Nature* 2012, **485**(7400):635.
15. Zhang J, Li JB, Liu BB, Chen J, Lu MZ: Expression and functional analysis of CDPK gene family in populus. *Forest Research* 2014.
16. Tuskan GA, Difazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, Putnam N, Ralph S, Rombauts S, Salamov A: The genome of black cottonwood, *Populus trichocarpa* (Torr.&Gray). *Science* 2006, **313**(5793):1596-1604.
17. Tissue and cell-type co-expression networks of transcription factors and wood component genes in *Populus trichocarpa*
18. Lobo I: Basic Local Alignment Search Tool (BLAST). *Journal of Molecular Biology* 2012, **215**(3):403-410.
19. Wang X, Shi X, Li Z, Zhu Q, Kong L, Tang W, Ge S, Luo J: Statistical inference of chromosomal homology based on gene colinearity and applications to *Arabidopsis* and rice. *BMC Bioinformatics* 2006, **7**(1):447.
20. Yang Z: PAML 4: phylogenetic analysis by maximum likelihood. *Molecular Biology & Evolution* 2007, **24**(8):1586-1591.
21. Robinson MD, Oshlack A: A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biology* 2010, **11**(3):1-9.
22. Evelyn C, Michele M, Daniel B, Vivian L, Emily D, John M, David B, Nicola H, Rodrigo L, Rolf A: The Gene Ontology Annotation (GOA) Database: sharing knowledge in Uniprot with Gene Ontology. *Nucleic Acids Research* 2004, **32**(Database issue):D262.
23. Yuki M, Masumi I, Shujiro O, Yoshizawa AC, Minoru K: KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Research* 2007, **35**(Web Server issue):182-185.
24. Kumar S, Stecher G, Tamura K: MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Molecular Biology & Evolution* 2016, **33**(7):1870.
25. Schnable JC, Springer NM, Freeling M: Differentiation of the maize subgenomes by genome dominance and both ancient and ongoing gene loss. *Proceedings of the National Academy of Sciences of the United States of America* 2011, **108**(10):4069-4074.
26. Jiao Y, Wickett NJ, Ayyampalayam S, Chanderbali AS, Landherr L, Ralph PE, Tomsho LP, Hu Y, Liang H, Soltis PS *et al*: Ancestral polyploidy in seed plants and angiosperms. *Nature* 2011, **473**(7345):97-100.

27. Jiao Y, Leebens-Mack J, Ayyampalayam S, Bowers JE, McKain MR, McNeal J, Rolf M, Ruzicka DR, Wafula E, Wickett NJ *et al*: **A genome triplication associated with early diversification of the core eudicots.** *Genome biology* 2012, **13**(1):R3.
28. Wang X, Wang J, Guo H, Lee T, Liu T, Jin D, Paterson AH: **Genome Alignment Spanning Major Poaceae Lineages Reveals Heterogeneous Evolutionary Rates and Alters Inferred Dates for Key Evolutionary Events.** *Molecular plant* 2015, **8**:14.
29. Chalhoub B, Denoeud F, Liu S, Parkin IA, Tang H, Wang X, Chiquet J, Belcram H, Tong C, Samans B *et al*: **Plant genetics. Early allopolyploid evolution in the post-Neolithic Brassica napus oilseed genome.** *Science* 2014, **345**(6199):950-953.
30. Wang J, Sun P, Li Y, Liu Y, Yu J, Ma X, Sun S, Yang N, Xia R, Lei T *et al*: **Hierarchically Aligning 10 Legume Genomes Establishes a Family-Level Genomics Platform.** *Plant physiology* 2017, **174**(1):284-300.
31. International Tomato Genome Sequencing Consortium: **The tomato genome sequence provides insights into fleshy fruit evolution.** *Nature* 2012, **485**(7400):635-641.
32. Wang J, Sun P, Li Y, Liu Y, Yang N, Yu J, Ma X, Sun S, Xia R, Liu X *et al*: **An overlooked paleo-tetraploidization in Cucurbitaceae** *Molecular and Biological Evolution* 2018, **35**(1):16-26.
33. Coate JE, Luciano AK, Seralathan V, Minchew KJ, Owens TG, Doyle JJ: **ANATOMICAL, BIOCHEMICAL, AND PHOTOSYNTHETIC RESPONSES TO RECENT ALLOPOLYPLOIDY IN GLYCINE DOLICHOCARPA (FABACEAE).** *American journal of botany* 2012, **99**(1):55-67.
34. Wang X, Wang H, Wang J, Sun R, Wu J, Liu S, Bai Y, Mun JH, Bancroft I, Cheng F *et al*: **The genome of the mesopolyploid crop species Brassica rapa.** *Nature genetics* 2011, **43**(10):1035-1039.
35. Liu S, Liu Y, Yang X, Tong C, Edwards D, Parkin IA, Zhao M, Ma J, Yu J, Huang S *et al*: **The Brassica oleracea genome reveals the asymmetrical evolution of polyploid genomes.** *Nature communications* 2014, **5**:3930.
36. Zhang H, Bian Y, Gou X, Zhu B, Xu C, Qi B, Li N, Rustgi S, Zhou H, Han F *et al*: **Persistent whole-chromosome aneuploidy is generally associated with nascent allohexaploid wheat.** *Proceedings of the National Academy of Sciences of the United States of America* 2013, **110**(9):3447-3452.
37. Zhao M, Meyers BC, Cai C, Xu W, Ma J: **Evolutionary patterns and coevolutionary consequences of MIRNA genes and microRNA targets triggered by multiple mechanisms of genomic duplications in soybean.** *The Plant cell* 2015, **27**(3):546-562.
38. Wang JP, Yu JG, Li J, Sun PC, Wang L, Yuan JQ, Meng FB, Sun SR, Li YX, Lei TY *et al*: **Two Likely Auto-Tetraploidization Events Shaped Kiwifruit Genome and Contributed to Establishment of the Actinidiaceae Family.** *iScience* 2018, **7**:230-240.

Tables

Table 1. Statistics of two sets of chromosome expression data after whole genome duplication of poplar

	Vv1	Vv2	Vv3	Vv4	Vv5	Vv6	Vv7	Vv8	Vv9	Vv10	Vv11	Vv12	Vv13	Vv14	Vv15	Vv16
Three-cell-type-1	0.681	0.061	0.641	0.692	0.588	0.048	0.362	0.574	0.428	0.106	0.902	0.726	0.931	0.118	0.413	0.793
Three-cell-type-2	0.707	0.106	0.932	0.786	0.503	0.287	0.112	0.625	0.363	0.051	0.841	0.966	0.731	0.072	0.050	0.711
Three-cell-type-3	0.680	0.051	0.565	0.414	0.418	0.340	0.195	0.584	0.405	0.108	0.971	0.915	0.130	0.099	0.179	0.725
Fiber-1	0.457	0.069	0.267	0.820	0.645	0.135	0.321	0.808	0.302	0.754	0.838	0.361	0.588	0.151	0.282	0.548
Fiber-2	0.672	0.113	0.555	0.764	0.838	0.296	0.952	0.673	0.515	0.271	0.899	0.479	0.375	0.117	0.033	0.774
Fiber-3	0.405	0.475	0.273	0.873	0.976	0.187	0.063	0.844	0.549	0.330	0.891	0.440	0.486	0.164	0.031	0.481
Vessel-1	0.404	0.088	0.793	0.181	0.426	0.746	0.532	0.795	0.584	0.067	0.370	0.706	0.424	0.227	0.781	0.640
Vessel-2	0.452	0.269	0.177	0.976	0.389	0.868	0.484	0.516	0.983	0.182	0.025	0.654	0.244	0.216	0.563	0.643
Vessel-3	0.129	0.300	0.630	0.342	0.196	0.829	0.324	0.668	0.993	0.018	0.476	0.168	0.315	0.121	0.875	0.614
Xylem-1	0.378	0.136	0.122	0.872	0.813	0.063	0.290	0.435	0.732	0.224	0.624	0.110	0.502	0.461	0.074	0.515
Xylem-2	0.399	0.145	0.131	0.402	0.971	0.203	0.227	0.690	0.628	0.119	0.631	0.326	0.864	0.498	0.080	0.659
Xylem-3	0.450	0.108	0.112	0.629	0.965	0.175	0.259	0.403	0.912	0.261	0.643	0.089	0.249	0.574	0.112	0.530
Phloem-1	0.564	0.242	0.651	0.399	0.223	0.860	0.120	0.562	0.681	0.385	0.293	0.389	0.143	0.904	0.041	0.803
Phloem-2	0.687	0.550	0.244	0.135	0.200	0.914	0.093	0.481	0.425	0.037	0.240	0.679	0.131	0.612	0.046	0.965
Phloem-3	0.620	0.203	0.777	0.444	0.238	0.785	0.106	0.503	0.697	0.400	0.295	0.513	0.162	0.715	0.047	0.666
Leaf-1	0.247	0.043	0.771	0.125	0.388	0.817	0.902	0.678	0.721	0.807	0.364	0.501	0.176	0.160	0.949	0.513
Leaf-2	0.281	0.066	0.905	0.117	0.507	0.719	0.984	0.326	0.547	0.729	0.357	0.481	0.210	0.262	0.887	0.651
Leaf-3	0.464	0.080	0.675	0.194	0.295	0.711	0.710	0.473	0.673	0.807	0.297	0.563	0.184	0.547	0.696	0.394
Shoot-1	0.815	0.144	0.248	0.676	0.248	0.326	0.844	0.115	0.821	0.838	0.637	0.255	0.096	0.835	0.473	0.517
Shoot-2	0.804	0.099	0.141	0.571	0.264	0.438	0.565	0.165	0.781	0.498	0.309	0.743	0.112	0.580	0.913	0.447
Shoot-3	0.758	0.125	0.186	0.572	0.293	0.509	0.387	0.218	0.996	0.804	0.141	0.691	0.119	0.674	0.887	0.672
Root-1	0.607	0.539	0.668	0.400	0.212	0.497	0.242	0.717	0.937	0.336	0.879	0.320	0.856	0.644	0.034	0.292
Root-2	0.642	0.468	0.630	0.393	0.218	0.518	0.205	0.712	0.958	0.319	0.888	0.344	0.885	0.631	0.024	0.484
Root-3	0.385	0.271	0.910	0.382	0.351	0.703	0.320	0.951	0.847	0.214	0.989	0.514	0.828	0.762	0.056	0.626
Average value	0.529	0.198	0.500	0.507	0.465	0.499	0.400	0.563	0.687	0.361	0.575	0.497	0.409	0.423	0.355	0.611

Note: Each row represents a different type of expression; each column represents two groups of poplar genes homologous to the grape chromosome, with the grape chromosome as an external group. The values in the table represent the results of paired-sample t test sig values for the two groups of poplar gene expression data.

Supplemental Figure Legend

Supplementary Figure 1. Dotplot of *Populus trichocarpa* and grape genome homologous genes. We distinguished the two subgenomes of poplar by referring to 19 chromosomes of grape and identified them with purple rectangle and green rectangle, respectively.

Figures

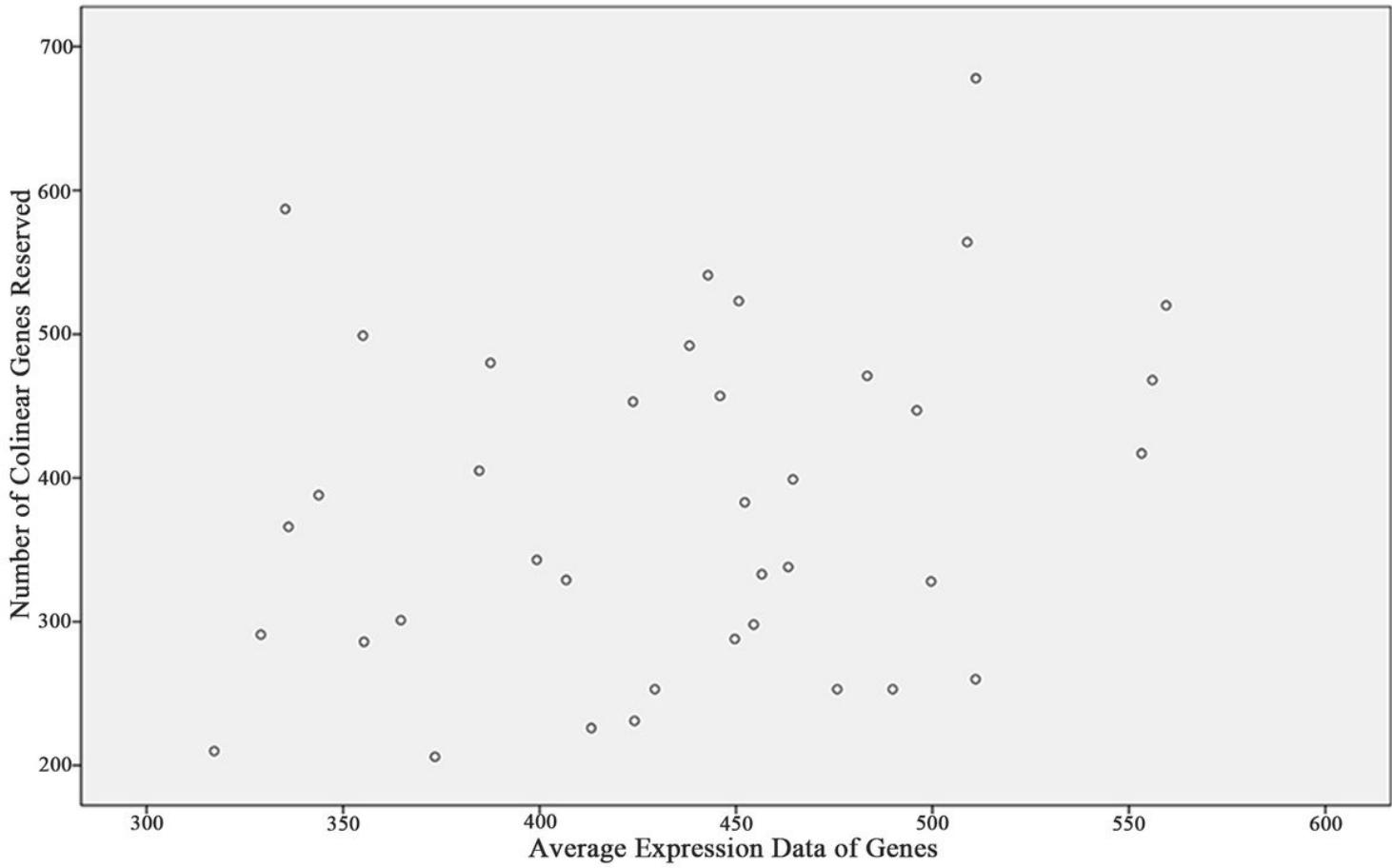


Figure 1

Spot Map of Colinear Gene Retention and Gene Expression in Poplar Subgenomes. The vertical axis represents the number of coplanar genes retained in the poplar subgenomes, while the horizontal axis represents the average gene expression of the subgenomes.

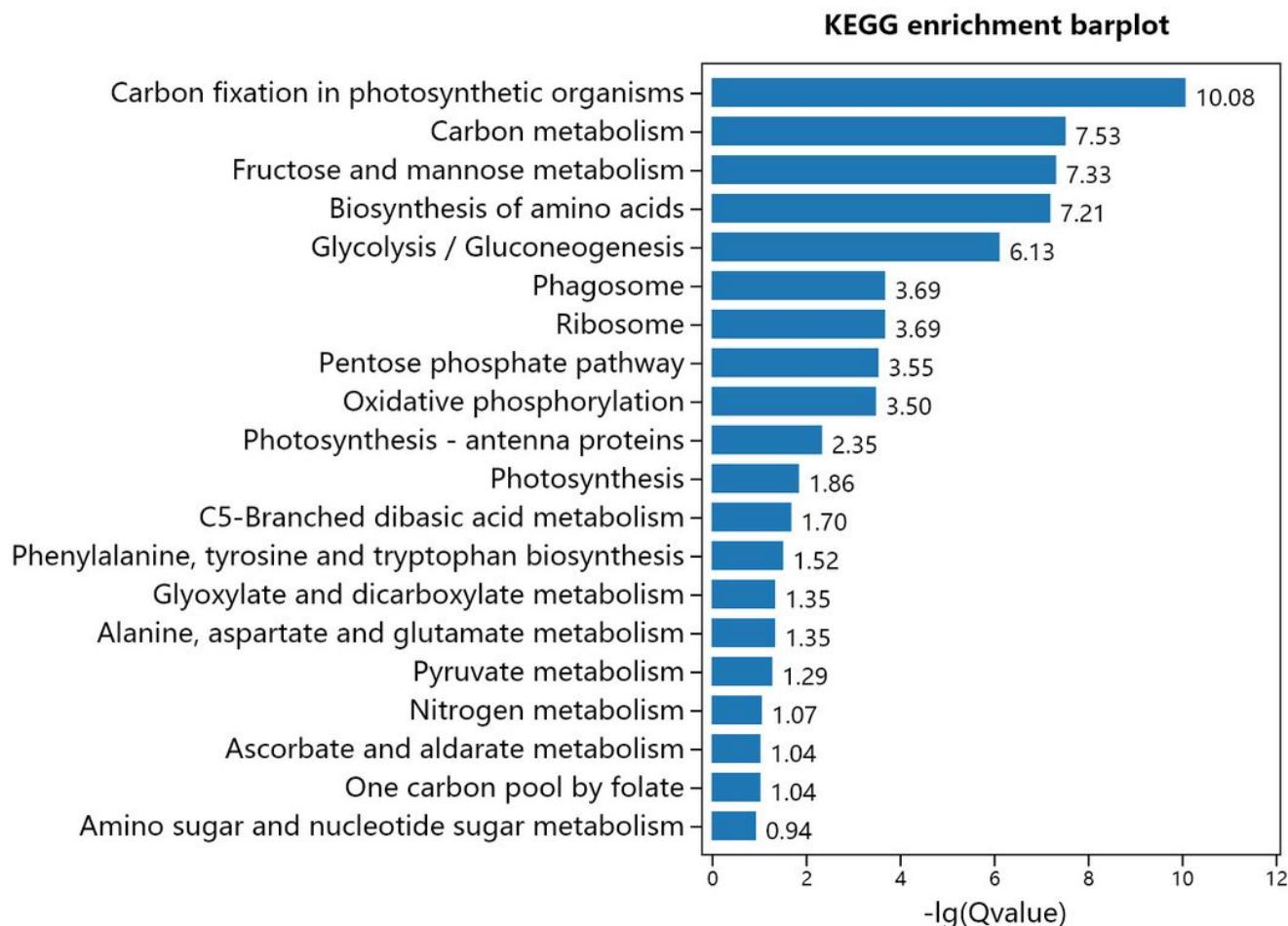


Figure 2

KEGG enrichment analysis of two sets of differentially expressed genes in Poplar. The vertical axis of the graph represents the different gene functions under the KEGG annotation, and the horizontal axis represents the value of $-\lg(\text{Pvalue})$. The larger the value, the more credible the enrichment of the function is.

KEGG pathway annotation

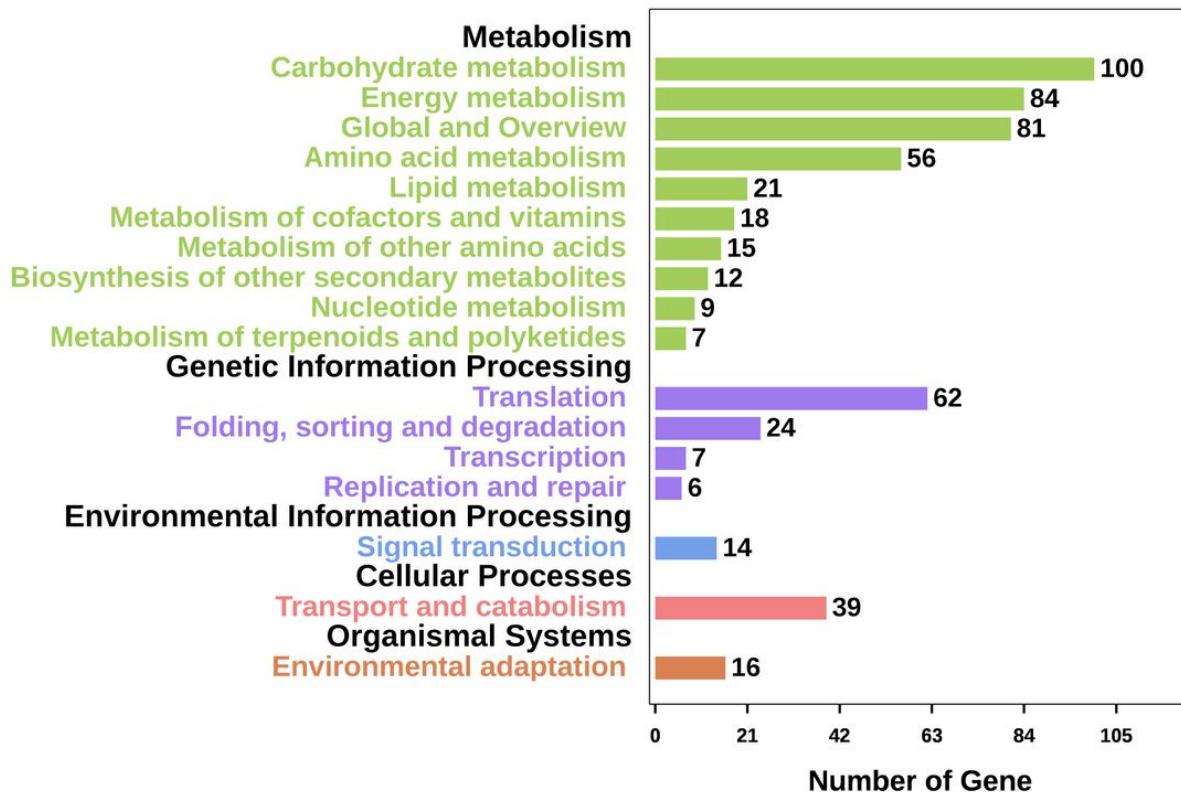


Figure 3

KEGG pathway annotation of two sets of differentially expressed genes in Poplar. The vertical axis is the main function under different pathways, and the horizontal axis is the number of genes enriched under this pathway. Different colors represent different main pathways.

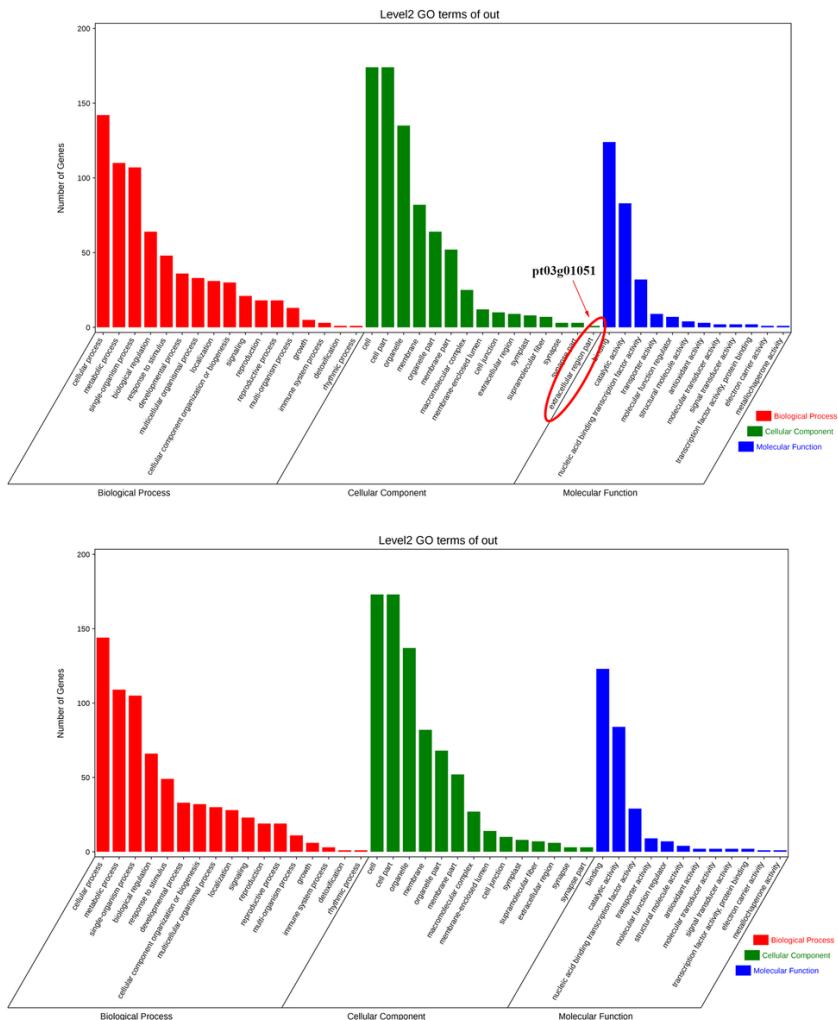


Figure 4

GO enrichment analysis of two sets of poplar gene collinear with grape chromosome 2. The horizontal axis in the graph represents different gene functions under GO annotation, and the longitudinal axis indicates the number of genes enriched by this function. Colors represent the three ontologies of GO annotation (cellular component, biological process, molecular function). The upper and lower parts represent the two sets of chromosomes PoplarA and PoplarB of poplar respectively, in which the different functions are marked with red ellipses.

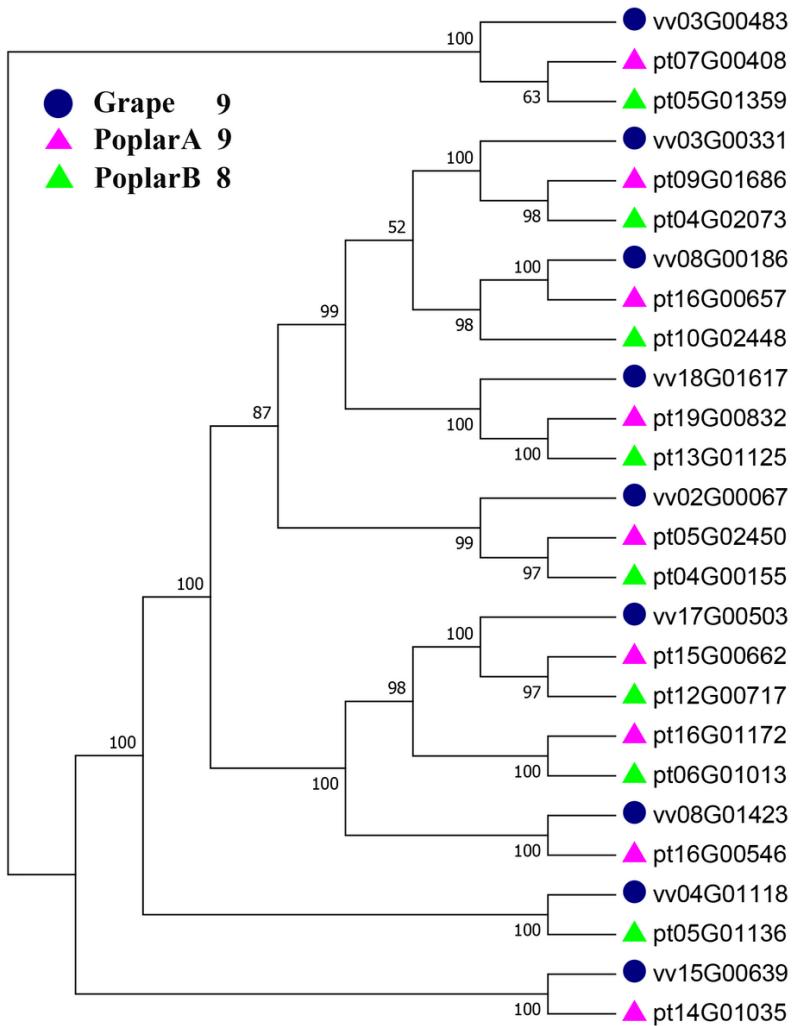


Figure 5

Phylogenetic tree of poplar and grape CDPK genes. CDPK protein sequences were aligned by Clustalw and the phylogenetic tree was constructed using Mega by the maximum likelihood method. Bootstrap value are based on 1000 replicates. Purple circle, pink triangle and green triangle respectively represent grape and two sets of poplar genes. The number on the branches is support value by bootstrapping.

Supplementary Files

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

- Supplementaryfig.1.jpg
- SupplementaryTable1.xls