Tracing the Transcriptomic Changes in Synthetic Trigenomic allohexaploids of *Brassica* Using an RNA-Seq Approach

Qin Zhao[1], Jun Zou[2], Jinling Meng[2], Shiyong Mei[3], Jianbo Wang[1*]

1 National Key Laboratory of Hybrid Rice, College of Life Sciences, Wuhan University, Wuhan, China, **2** National Key Laboratory of Crop Genetic Improvement, Huazhong Agricultural University, Wuhan, China, **3** Hubei Academy of Agricultural Science, Wuhan, China

Abstract

Polyploidization has played an important role in plant evolution and speciation, and newly formed allopolyploids have experienced rapid transcriptomic changes. Here, we compared the transcriptomic differences between a synthetic *Brassica* allohexaploid and its parents using a high-throughput RNA-Seq method. A total of 35,644,409 sequence reads were generated, and 32,642 genes were aligned from the data. Totals of 29,260, 29,060, and 29,697 genes were identified in *Brassica rapa*, *Brassica carinata*, and *Brassica* allohexaploid, respectively. We compared 7,397 differentially expressed genes (DEGs) between *Brassica* hexaploid and its parents, as well as 2,545 nonadditive genes of *Brassica* hexaploid. We hypothesized that the higher ploidy level as well as secondary polyploidy might have influenced these changes. The majority of the 3,184 DEGs between *Brassica* hexaploid and its paternal parent, *B. rapa*, were involved in the biosynthesis of secondary metabolites, plant–pathogen interactions, photosynthesis, and circadian rhythm. Among the 2,233 DEGs between *Brassica* hexaploid and its maternal parent, *B. carinata*, several played roles in plant–pathogen interactions, plant hormone signal transduction, ribosomes, limonene and pinene degradation, photosynthesis, and biosynthesis of secondary metabolites. There were more significant differences in gene expression between the allohexaploid and its paternal parent than between it and its maternal parent, possibly partly because of cytoplasmic and maternal effects. Specific functional categories were enriched among the 2,545 nonadditive genes of *Brassica* hexaploid compared with the additive genes; the categories included response to stimulus, immune system process, cellular process, metabolic process, rhythmic process, and pigmentation. Many transcription factor genes, methyltransferases, and methylation genes showed differential expression between *Brassica* hexaploid and its parents. Our results demonstrate that the *Brassica* allohexaploid can generate extensive transcriptomic diversity compared with its parents. These changes may contribute to the normal growth and reproduction of allohexaploids.

Introduction

Polyploidization is an important evolutionary process in eukaryotes, and it is believed that all angiosperms underwent a whole genome duplication event during their evolution [1]. Plant genome sequencing analysis has revealed that species once considered diploids, such as *Arabidopsis thaliana*, *Vinis vitifera*, *Oryza sativa*, and *Populus trichocarpa*, are possibly paleopolyploids [2–4] that experienced at least one whole genome duplication event followed by a process of genomic reorganization known as diploidization [5].

Allopolyploids originate through hybridization between different species followed by chromosome duplication. They

have contributed to the origin of species and have driven the evolution of angiosperms [6]. The merging of genomes previously adapted to different environments allows allopolyploids to adapt to a wider range of environmental conditions [7]. Some important crops such as wheat, oat, cotton, coffee, and rapeseeds are allopolyploids [8]. The common occurrence of allopolyploids in nature suggests that they have a selective advantage, possibly due in part to heterozygosity and gene redundancy [9,10]. However, increasing gene and genome dosages in allopolyploids also leads to massive gene loss and genome rearrangement. Molecular studies have revealed dynamic and pervasive changes in the polyploid genome and transcriptome that vary

among different allopolyploids; these include DNA sequence elimination [11,12], genome rearrangement [13,14], transposon activation [15,16], gene silencing [17], and alterations of gene expression [18,19].

Recent studies have indicated that the transcriptomic changes in allopolyploids may be an adaptive mechanism that facilitates the establishment of gene expression programs and the evolution of a stable species [20]. Changes in gene expression result in the subfunctionalization and neofunctionalization of genes. Subfunctionalization can occur when a homologous gene originating from one parental genome is silenced in specific organs, while the other is silenced in other organs. This is similar to homologs of the alcohol dehydrogenase gene *Adh* in *Gossypium hirsutum* [21]. The duplicated *BnPOX* in the allotetraploid *Brassica napus* retained its function by subfunctionalization of the homologs in different organs, as well as in response to biotic stress [22]. It has been suggested that subfunctionalization is a transition state to neofunctionalization [23]. Transcriptional subfunctionalization and neofunctionalization were both detected in allopolyploid cotton [24]. Changes in gene expression may vary in different allopolyploids, and thus further studies are required.

Molecular markers and microarray data have been applied to analyze genomic and transcriptomic changes in allopolyploids [13,25,26]. The sequencing-based method RNA-Seq allows the entire transcriptome to be surveyed in a high-throughput and quantitative manner. Compared with microarrays and other sequence-based approaches using DNA libraries, RNA-Seq has a low background signal, a large dynamic range of expression levels, more accurate quantification, and high levels of reproducibility [27]. Massively parallel sequencing of DNA using pyrosequencing technology has been applied to obtain an unbiased representation of *Arabidopsis* transcripts [28]. For *Medicago truncatula*, a single 454 Life Science GS20 sequencing approach using a normalized cDNA library has been used to identify novel genes and effectively detect transcript expression [29]. RNA-Seq has provided a powerful approach for determining the expression levels of transcripts and gene function annotation in rice and soybean [30–32]. RNA-Seq has also been applied to explore transcriptomic changes during berry development in grapes, to investigate the transcriptional response of sugar beets to vernalization and gibberellin treatment, and to identify genes involved in cucumber sex determination [16,33,34].

Many studies have used artificially synthesized polyploid materials to determine how duplicated genes and genomes maintain and lose functions during the early stages of polyploidization [35]. Synthesized polyploids can be compared with their direct parental lines, providing an advantage over natural polyploids. The genus *Brassica* is commonly used as a model system in which to perform molecular and phenotypic characterizations of polyploids. Also, the allotetraploid *B. napus* is often synthesized for polyploid studies [26,36]. The synthesized hexaploid *Brassica* is relatively rare, but has played an important role in allopolyploid studies [37]. However, little information is available for artificially synthesized polyploids after the first generation, possibly due to instability.

Here, we used a relatively stable trigenomic *Brassica* allohexaploid material and its parents as an experimental system [38].

In this study, we applied RNA-Seq to identify differential gene expression and compared gene function between the trigenomic *Brassica* allohexaploid (BBCCAA, 2n = 54) and its parents. More differences in gene expression were observed between hexaploid *Brassica* and the paternal parent *B. rapa* (AA, 2n = 20) than between it and its maternal parent *B. carinata* (BBCC, 2n = 34). The differentially expressed genes were involved in a wide range of development and stress resistance processes.

Materials and Methods

Plant materials and RNA extraction

A trigenomic *Brassica* allohexaploid (BBCCAA genome) and its parents *B. carinata* (BBCC) and *B. rapa* (AA) were used in this study. The *B. carinata* (accession number CGN03953 in the Center for Genetic Resources, Wageningen, the Netherlands), and Chinese *B. rapa* (cultivar name, Dongkoutian) were grown and self-pollinated in the field of Huazhong Agricultural University. The trigenomic *Brassica* allohexaploid was generated by inter-specific hybridization followed with chromosome doubling, by using *B. carinata* as the maternal parent and *B. rapa* as the paternal parent in the year of 2004 [38]. The synthesized allohexaploid (code as C14) used in this study was in the fourth generation. After chromosome identification of the synthesized allohexaploid, it was confirmed to be euploid (2n=54). The plant materials were grown in a greenhouse at Hubei Academy of Agricultural Science under natural conditions of light and temperature (Wuhan, China). Young leaves from five replicate individuals of both parents and the allohexaploid were collected for RNA extraction when the materials have grown for three months.

For Illumina sequencing, total RNA was isolated using TRIzol reagent according to the manufacturer's protocal (Invitrogen, Burlington, ON, Canada) and treated with RNase-free DNase I (Fermentas, Canada). RNA quality was confirmed using a 1% agarose gel with the RNA 6000 Nano Assay Kit and Agilent 2100 Bioanalyzer. The total RNA was stored at -80°C for later use.

cDNA library construction and sequencing

Poly (A)-containing mRNA molecules were purified from total RNA using poly (T) oligo-attached magnetic beads. The mRNA was fragmented into small pieces (59±1bp) by treatment with divalent cations at elevated temperature. Then the double-stranded cDNA was synthesized using the SuperScript Double-Stranded cDNA Synthesized kit (Invitrogen, Camarillo, CA, USA) and random hexamer primers (Illumina, San Diego, CA, USA). The double-stranded cDNA was further subjected to end-repair and phosphorylation using T4 DNA polymerase, Klenow DNA polymerase and T4 PNK (Fermentas, canada). These repaired cDNA fragments were 3' adenylated using Klenow 3' to 5' exo-polymerase and ligated Illumina Paired-end adapters to the ends. The required fragments were separated by size on an agarose gel and excised from the gel. PCR was

performed to enrich the purified cDNA fragments. After validation with a Bioanalyzer Chip DNA 1000 series II (Agilent, Santa Clara, CA, USA), the cDNA library was sequenced on an Illumina HiSeq2000 sequencing platform. Files containing reads sequence and quality scores were deposited in the Gene Expression Omnibus of National Center for Biotechnology Information with accession numbers GSE46299.

Mapping of reads and expression analysis of genes

The raw Illumina sequencing files were base called into sequence data. We cleaned the raw reads by removing adapter sequences and low-quality sequences (reads has more than 50% of low-quality bases with a quality value ≤ 5). The *B. rapa* genome and gene information were downloaded from the *Brassica* database (BRAD, http://brassica.org/brad/) [39]. We used SOAPaligner/soap2 software to map the reads to the *B. rapa* reference genome with a permission of two base mismatches.

For gene expression analysis, the raw gene expression counts were measured by quantifying the number of reads that were mapped to the *B. rapa* genome sequences using ERANGE 4.0 software (http://woldlab.caltech.edu/gitweb/). ERANGE reports the number of mapped reads per kilobase of exon per million mapped reads, measuring the transcriptional activity for each gene [40].

Gene annotation

To annotate the detected genes, a BLASTx search against the NCBI non-redundant protein (Nr) database (http://www.ncbi.nlm.nih.gov/) and the Swiss-Prot protein database was performed with an E-value threshold of less than 10^{-5}. With Nr annotation, Blast2GO [41] was used to obtain GO annotation of genes according to molecular function, biological process and cellular component ontologies (http://www.geneontology.org/). WEGO [42] was used for GO functional classification of all genes and to plot the distribution of gene functions. Pathway assignments were carried out based on the KEGG database [43].

The putative TFs in these genes were got by querying *B. rapa* Transcription Factors in the Plant Transcription Factor Database (http://planttfdb.cbi.edu.cn/).

Differential gene expression analysis

Significantly differential expressed genes were determined using R package DEGseq [44]. We defined genes with fold change ≥ 2 and FDR ≤ 0.001 as differentially expressed genes (DEGs). GO terms enrichment was performed to the DEGs. To group DEGs with similar expression patterns, a hierarchical clustering was generated using the expression values from each library. The analysis was conducted using MATLAB 7.8 software with Pearson correlation as the distance measure. The cluster tree concluded distinct clusters which contained genes with a unique expression profile by visual inspection.

Table 1. Number of reads matched to *B. rapa* genome.

	B. rapa	B. carinata	Brassica hexaploid
Total reads	12,216,743	11,859,888	11,567,778
Mapped reads	9,359,218 (76.61%)	6,360,133 (53.63%)	7,026,728 (60.74%)
Unique match	8,594,083 (70.35%)	5,479,383 (46.20%)	6,267,830 (54.18%)
Multi-position match	765,135 (6.26%)	880,750 (7.43%)	758,898 (6.56%)
Unmapped reads	2,857,525 (23.39%)	5,499,755 (46.37%)	4,541,050 (39.26%)

Results

Illumina sequencing and the data analysis

We previously performed a *B. carinata* × *B. rapa* cross to double the genome, with *B. carinata* as the maternal parent and *B. rapa* as the paternal parent [38]. We used the RNA-Seq method to compare differences in gene expression between *Brassica* hexaploid and its parents. Using the Illumina HiSeq 2000 platform, a total of 35,644,409 sequence reads (59 bp in length) were generated from cDNA libraries derived from *B. rapa*, *B. carinata*, and *Brassica* hexaploid. Each library generated more than 10 million reads, reaching the saturation level of gene identification. SOAPaligner/soap2 software was used to map the reads to the *B. rapa* reference genome (http://brassica.org/brad/) with a tolerance of two base mismatches. Reads that mapped to a unique sequence were the most critical subset of the RNA-Seq libraries, as they can explicitly identify a gene. In total, 70.35%, 46.20%, and 54.18% of reads uniquely matched to a genomic location in the *B. rapa*, *B. carinata*, and *Brassica* hexaploid libraries, while unmapped reads accounted for 23.39%, 46.37%, and 39.26%, respectively (Table 1). It is possible that *B. carinata* had the smallest percentage of mapped reads because of genome differentiation between *B. carinata* (BBCC) and *B. rapa* (AA, the reference genome for gene annotation). Meanwhile, the hexaploid received a middle-parent number. Digital gene expression analysis was performed for the uniquely mapped reads. To compare transcript levels within and between samples, the raw read counts were normalized using a variation of the reads/Kb/million (RPKM) method. The RPKM method can avoid biases caused by different gene exon sizes and the different number of read sequences obtained from the three libraries.

Our RNA-Seq data revealed 32,642 genes in *Brassica* hexaploid and its parents, accounting for 79.28% of the inferred genes in the *B. rapa* genome [39]. To confirm whether the number of detected genes increased proportionally with the sequencing amount (total tag number), a saturation analysis was performed. The saturation trend (Figure 1) showed that the number of detected genes stopped increasing when the number of reads reached 2 million. A total of 70.95% of the detected genes had at least 30 matching reads when including
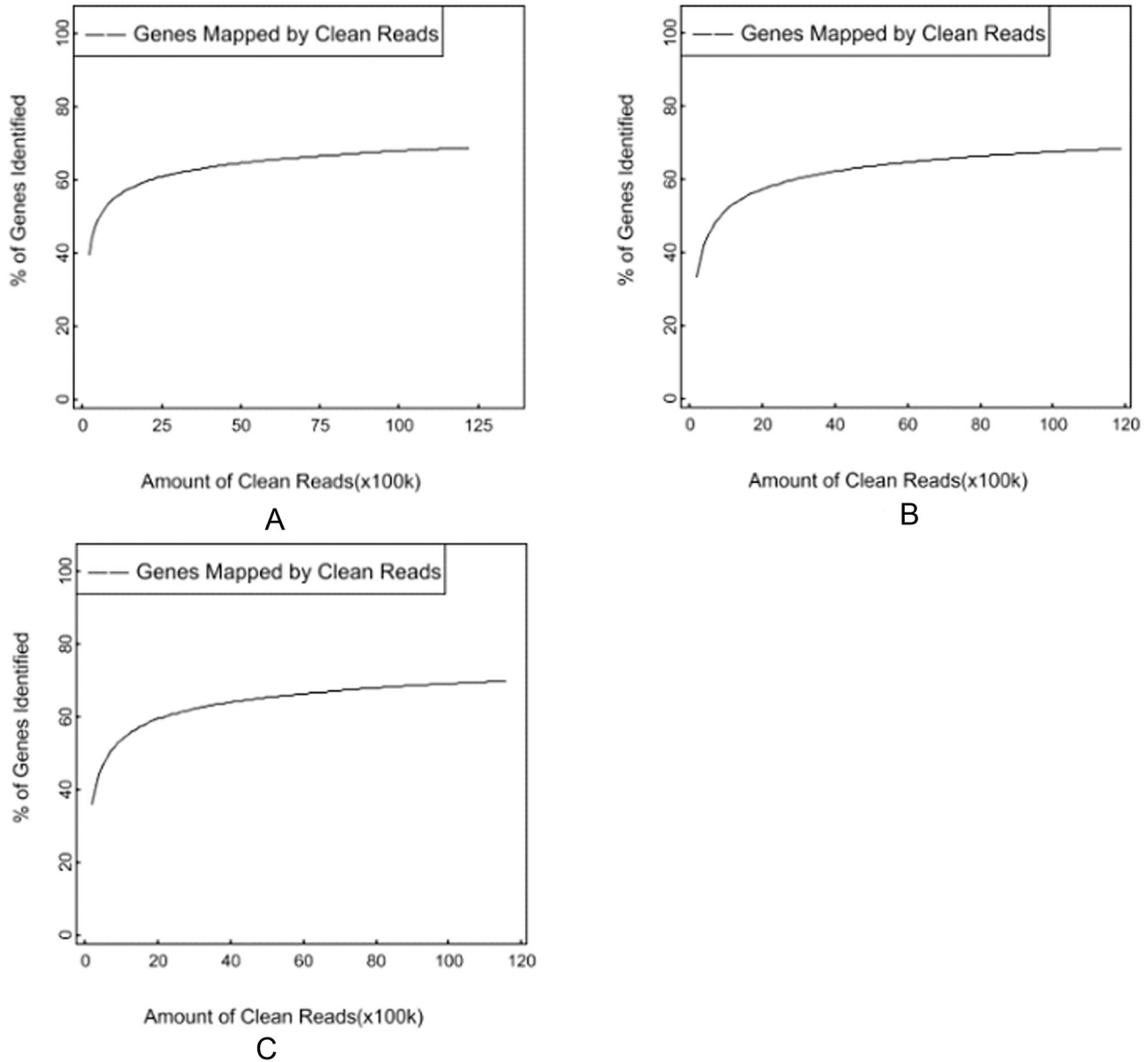
Figure 1. The saturation analysis of genes detected by sequencing reads. The gene saturation analysis are taken in the libraries of *B. rapa* (A), *B. carinata* (B), and *Brassica* hexaploid (C) respectively. The curve indicated that the number of detected genes almost ceases to increase when the number of reads reaches 2 million.

doi: 10.1371/journal.pone.0068883.g001

all three libraries. The genes matched to reads had different percentages of their sequences covered by the reads, possibly because of the heterogeneity and redundancy of mRNAs. Genes having more than 90% of their sequence covered by reads constituted the most abundant category, accounting for 38%, 15%, and 26% of the matched genes in *B. rapa*, *B. carinata*, and *Brassica* hexaploid, respectively. The second most abundant category was genes with 80% to 90% sequence coverage. The percentages of matched genes were similar among the other eight categories, established according to increments of 10% (Figure 2).

Genes expressed in *Brassica* hexaploid and its parents, and their functional analysis

The 32,642 genes expressed in *Brassica* hexaploid and its parents had lengths of 100 to ≥ 2,000 bp. Genes of 500–1,000 bp were the most abundant, accounting for 30.71% of the total genes, and genes with lengths between 1,000 and 1,500 bp accounted for 25.41%. Genes with lengths of 100–500 bp,
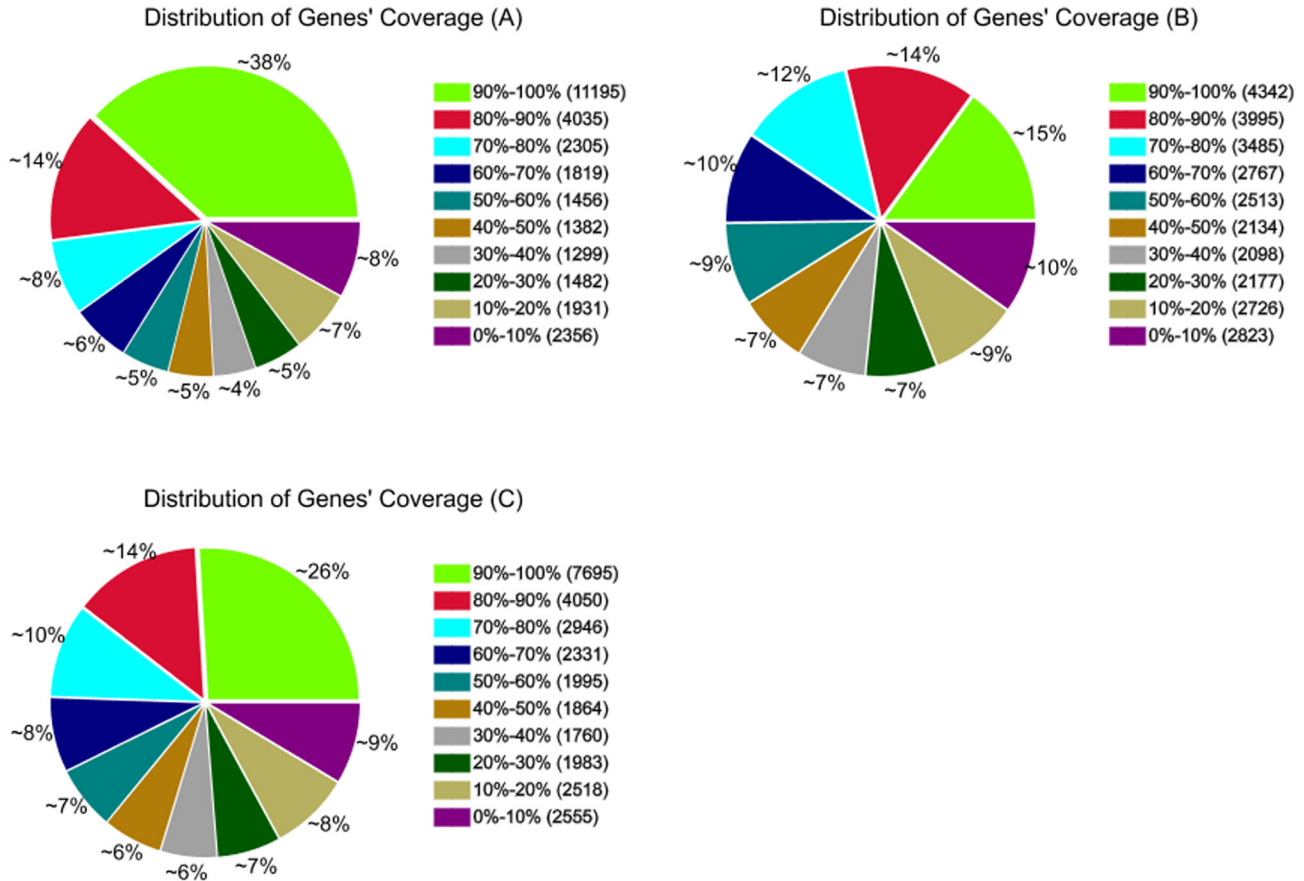
Distribution of Genes' Coverage (A)

90%-100%	(11195)
80%-90%	(4035)
70%-80%	(2305)
60%-70%	(1819)
50%-60%	(1456)
40%-50%	(1382)
30%-40%	(1299)
20%-30%	(1482)
10%-20%	(1931)
0%-10%	(2356)

Distribution of Genes' Coverage (B)

90%-100%	(4342)
80%-90%	(3995)
70%-80%	(3485)
60%-70%	(2767)
50%-60%	(2513)
40%-50%	(2134)
30%-40%	(2098)
20%-30%	(2177)
10%-20%	(2726)
0%-10%	(2823)

Distribution of Genes' Coverage (C)

90%-100%	(7695)
80%-90%	(4050)
70%-80%	(2946)
60%-70%	(2331)
50%-60%	(1995)
40%-50%	(1864)
30%-40%	(1760)
20%-30%	(1983)
10%-20%	(2518)
0%-10%	(2555)

Figure 2. The different read coverage of genes in three libraries. The genes with different read coverage were indicated by different colours. The capital letters A, B, and C in brackets refer to libraries of *B. rapa*, *B. carinata*, and *Brassica* hexaploid respectively.

doi: 10.1371/journal.pone.0068883.g002

1,500–2,000 bp, and ≥ 2,000 bp represented 15.46%, 13.62%, 14.80% of the total genes, respectively (Table 2).

Similar numbers of genes were detected in *B. rapa*, *B. carinata*, and the allohexaploid (Figure 3). Of the 25,953 genes shared by *Brassica* hexaploid and its parents, 628 were co-expressed in *B. rapa* and *B. carinata*, 1,552 were co-expressed in *B. carinata* and *Brassica* hexaploid, and 1,289 were co-expressed in *B. rapa* and *Brassica* hexaploid. However, 1,390 (*B. rapa*), 927 (*B. carinata*), and 903 (*Brassica* hexaploid) genes were specifically expressed, respectively. Among the 25,953 genes expressed in all three *Brassica* and the 3,469 genes co-expressed in two, many showed variable gene expression levels between *Brassica* hexaploid and its parents.

To further explore the functional significance of all detected genes and the genes specifically expressed in each library, we assigned the gene sequences to gene ontology (GO) functional classes (http://www.geneontology.org/) using Blast2GO (version 2.3.5) (http://www.blast2go.org/). A total of 32,642 genes were detected, and 24,658 (75.54%) had at least one GO annotation based on sequence similarity. We categorized these 24,658 genes according to the secondary classification

Table 2. Gene sequence length distribution.

Gene length (bp)	Total number	Percentage
100-500	5048	15.46%
500-1000	10024	30.71%
1000-1500	8293	25.41%
1500-2000	4446	13.62%
≥2000	4831	14.80%
Total	32642	100%

of GO terms, assigning genes to 46 functional groups of the three main GO classification categories (Figure 4). Genes were classified into 26 functional groups of the biological process GO category, 11 functional groups of the molecular function category, and nine functional groups of the cellular component category. The terms cellular process (GO:0009987) and metabolic process (GO:0008152), which included 12,537 and 12,007 genes, respectively, were dominant in biological process category. In the molecular function and cellular component categories, the terms binding (GO:0005488) and
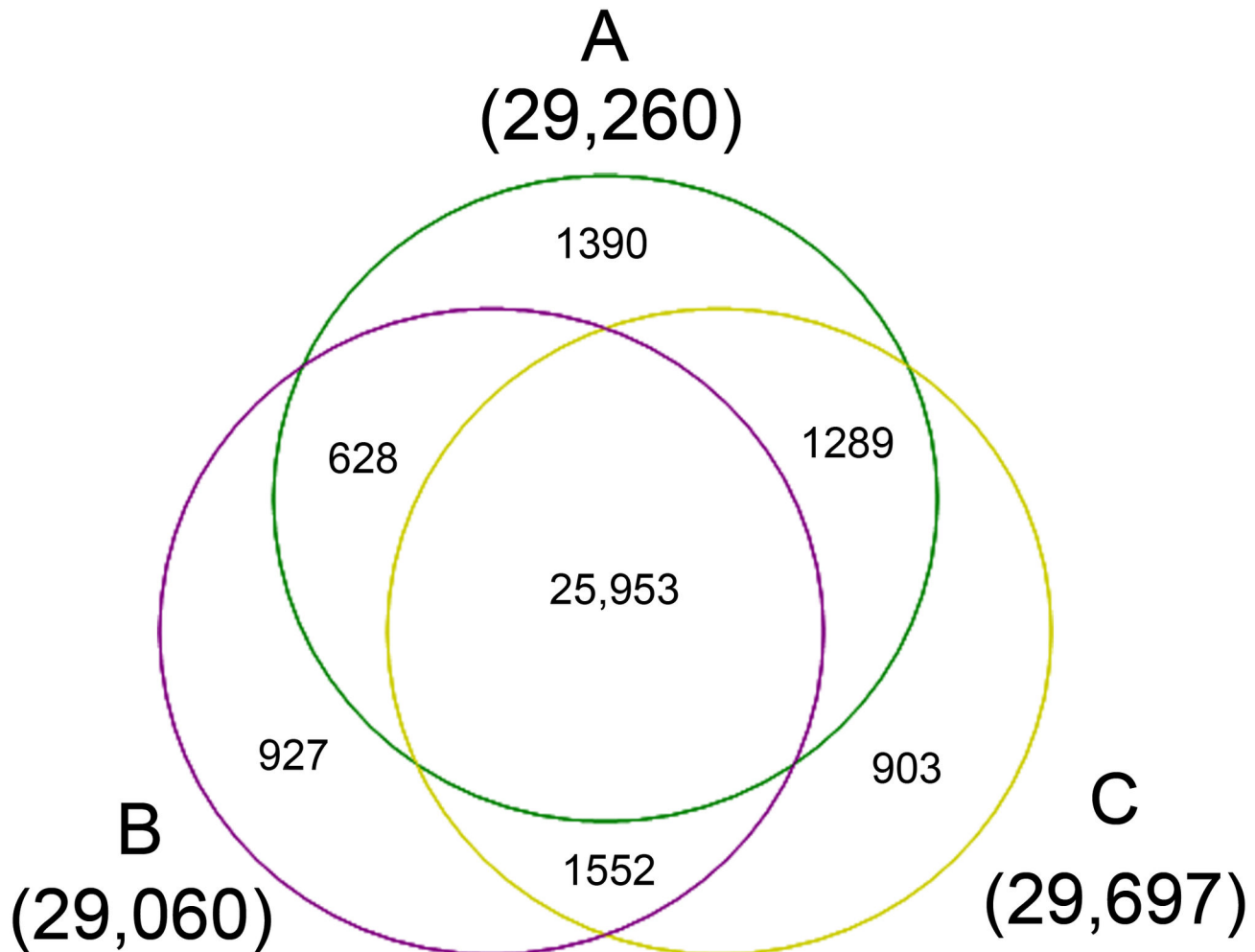
Figure 3. Venn diagram showing genes expressed in *Brassica* hexaploid and its parents. A, *B. rapa*; B, *B. carinata*; C, *Brassica* hexaploid.

cell (GO:0005623) were most significantly enriched, with 18,115 and 10,922 genes, respectively. There was also an enrichment for the terms response to stimulus (GO:0050896), catalytic activity (GO:0003824), and cell part (GO:0044464) in the three main categories. Among the genes specifically expressed in *B. rapa* (1,390 genes), *B. carinata* (927 genes), and *Brassica* hexaploid (903 genes), we observed a similarly high distribution of genes encoding proteins involved in metal ion binding and antioxidation, with enrichment of the GO terms transition metal ion binding, metal ion binding, binding, and oxidoreductase activity.

Pathway enrichment analysis was performed on all expressed transcripts by mapping them to reference canonical pathways in the Kyoto Encyclopedia of Genes and Genomes (KEGG) (http://www.genome.ad.jp/kegg/). A total of 21,092 genes were mapped to 125 KEGG pathways. Pathways with the highest numbers of genes included metabolic pathways (4,529 genes), biosynthesis of secondary metabolites (2,547

genes), plant–pathogen interactions (1,867 genes), and plant hormone signal transduction (1,684 genes) (Table S1). These pathways are involved in major growth processes and stress reactions in plants.

Differentially expressed genes and functional analysis

Differentially expressed genes (DEGs) between the *Brassica* hexaploid and its parents were identified using a likelihood ratio test [45] to compare the gene RPKM values. Applying the algorithm test built by Audic et al. [46], we considered significantly altered genes to be those with expression changes of at least two-fold and false discovery rates (FDR) of no more than 0.001. As a result, 7,397 genes showed significant differential expression between *Brassica* hexaploid and its parents. Comparing *Brassica* hexaploid and *B. rapa*, 2,670 genes were upregulated and 2,578 genes were downregulated (Figure 5 and Table S2). Compared with *B. carinata*, *Brassica* hexaploid had 2,274 upregulated genes and 1,370
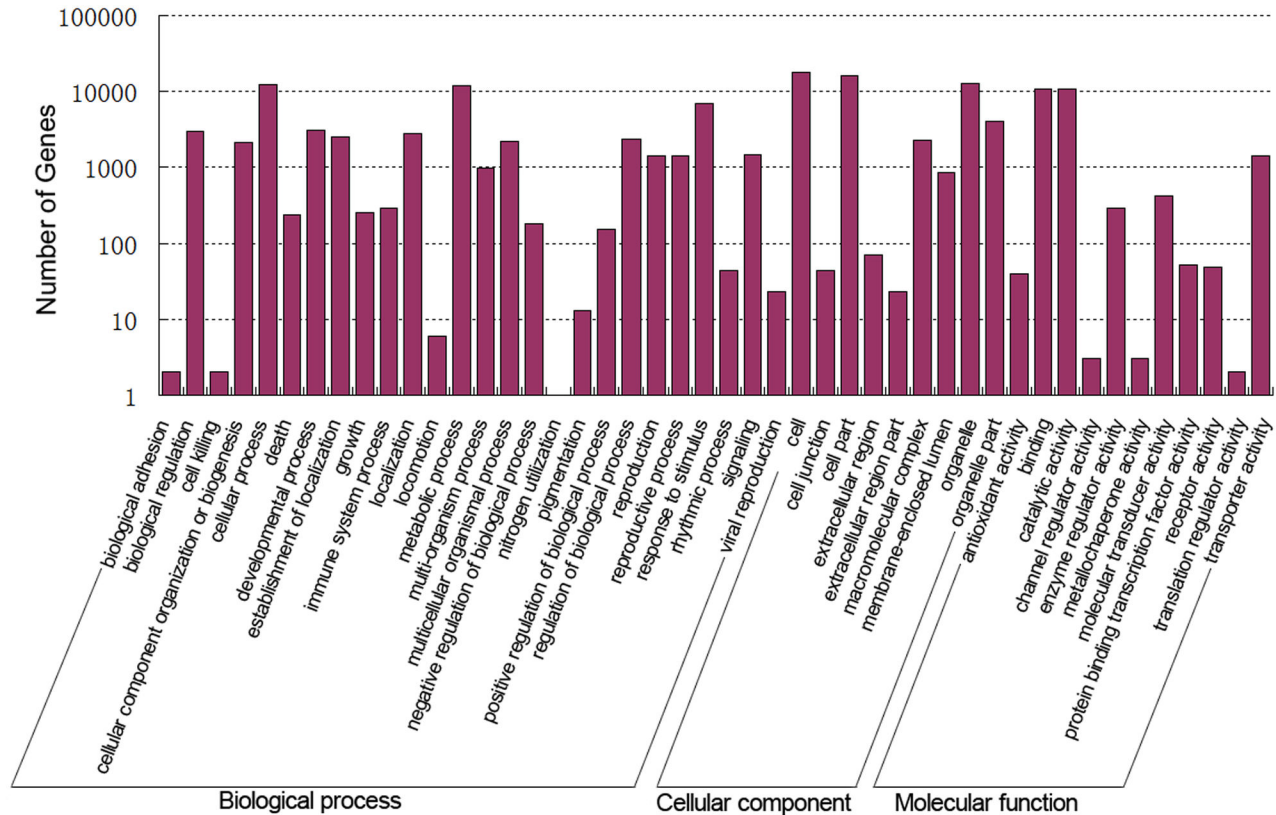
Figure 4. Gene ontology (GO) annotations of all detected genes. The histogram shows the result of classifying 24658 genes to the secondary classification of GO terms. The y-axis indicates the number of genes in a functional term.

doi: 10.1371/journal.pone.0068883.g004

downregulated genes (Figure 5 and Table S3). The total value based on the fold change of DEGs between *Brassica* hexaploid and *B. rapa* was 15,640.92. Between *Brassica* hexaploid and *B. carinata*, the value was 10,371.10. These values suggest that there were more DEGs with larger differences in expression between *Brassica* hexaploid and *B. rapa*.

Using GO term enrichment analysis, we categorized the DEGs between *Brassica* hexaploid and *B. rapa* and between *Brassica* hexaploid and *B. carinata* according to the secondary classification of GO terms. There were genes belonging to eight, nine, and 20 functional groups in the three main GO classification categories: molecular function, cellular component, and biological process, respectively (Figure 6). The terms binding (GO:0005488) and metabolic process (GO: 0008152) were dominant in the molecular function and biological process categories. In the cellular components category, the terms cell (GO:0005623) and cell part (GO: 0044464) were equally common and occurred more frequently than the other functional groups in this category. The catalytic activity (GO:0003824), transporter activity (GO:0005215), organelle (GO:0043226), cellular process (GO:0009987), and response to stimulus (GO: 0050896) functional groups were also significantly enriched in DEGs between *Brassica* hexaploid and its parents.

Mapping of the DEGs to KEGG pathways revealed significant enrichment of specific pathways compared with the distribution of the whole transcriptome. Among the DEGs between *Brassica* hexaploid and its paternal parent, 3,184 had a KEGG pathway annotation. Pathways related to secondary metabolite biosynthesis, amino acids, nitrogen, sulfur, lipid, vitamin B6 metabolism, plant–pathogen interactions, photosynthesis, and circadian rhythm were overrepresented. Between *Brassica* hexaploid and its maternal parent, 2,233 DEGs mapped to KEGG pathways, with enrichment of pathways regarding plant–pathogen interactions, plant hormone signal transduction, ribosomes, limonene and pinene degradation, photosynthesis, and the biosynthesis of secondary metabolites such as stilbenoid, diarylheptanoid, gingerol, glucosinolate, flavone, and flavonol. Thus, the DEGs between *Brassica* hexaploid and its parents were involved in a wide range of plant physiological processes that may be essential for the differences in morphology and physiology between *Brassica* hexaploid and its parents.

Clustering of differentially expressed genes

We generated a hierarchical clustering of the 7,397 DEGs to examine the similarity and diversity of expression profiles. The RPKM value was transformed by dividing the value of a gene in
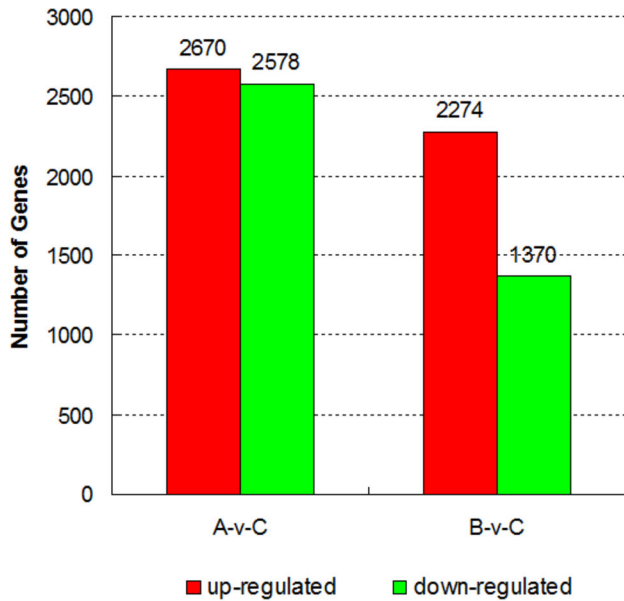
Figure 5. Differentially expressed genes detected between *Brassica* hexaploid and its parents. The number of up-regulated and down-regulated genes between *Brassica* hexaploid (C) and *B. rapa* (A), *Brassica* hexaploid (C) and *B. carinata* (B) are revealed.

one library by the control library value, and this was then expressed as the binary logarithm. Figure 7 shows the results as displayed by Java Treeview. The transformed values are represented by different colors, with red representing positive values and green representing negative values.

We analyzed the associations among DEGs based on the correlations and differences among their expression patterns as revealed by the hierarchical clustering. We classified 1,392 genes with significant differential expression between *Brassica* hexaploid and *B. rapa* and between *Brassica* hexaploid and *B. carinata* into four expression patterns using two-dimensional hierarchical clustering. Cluster 1 had 319 genes; cluster 2 had 218; cluster 3 had 187; and cluster 4 had 668 (Figure 8 and Table S4). Cluster 4, the most abundant group, was composed of genes with the highest expression in *Brassica* hexaploid. As the second most abundant group, cluster 1 contained genes negatively modulated in *Brassica* hexaploid compared with its parents. Genes in cluster 2 were downregulated in *Brassica* hexaploid compared with its paternal parent *B. rapa*, and were upregulated compared with its maternal parent *B. carinata*. Conversely, cluster 3 included genes upregulated in *Brassica* hexaploid compared with *B. rapa* and downregulated compared with *B. carinata*.

We also categorized 7,397 DEGs showing differential expression between *Brassica* hexaploid and *B. rapa* or *Brassica* hexaploid and *B. carinata* into four groups (Table S5). Genes in groups 1, 2, 3, and 4 had the same expression profiles as genes in clusters 1, 2, 3, and 4. There were 2,016, 1,701, 1,327, and 2,353 genes in groups 1, 2, 3, and 4,

respectively. We calculated the functional category abundance in the four groups and identified differences in gene distribution. Excluding genes encoding proteins with functions in adenyl ribonucleotide binding, nucleic acid binding, transition metal ion binding, protein binding, structural molecule activity, and protein kinase activity (which were enriched in all four groups), there were more genes involved in molecular transducer activity, hydrolase activity, and protein phosphatase regulator activity in group 1, and more genes in the functional categories of metal ion binding, transcription regulator activity, carboxylesterase, and DNA binding in group 2. For group 3, cofactor binding, transferase activity, endopeptidase activity, RNA binding, and guanyl ribonucleotide binding functions were common. Genes categorized as carbohydrate binding, phosphatase activity, oxidoreductase activity, catalytic activity, and kinase activity were more enriched in group 4 than in the other three groups.

Nonadditive genes expressed in *Brassica* hexaploid

We compared the expression levels of genes in *Brassica* hexaploid with the midparent value (MPV). Genes showing at least a two-fold change between *Brassica* hexaploid and MPV (with FDRs of no more than 0.001) were considered nonadditive genes; all others were considered additive genes. A total of 2,545 genes, accounting for 7.797% of the total detected genes, were nonadditively expressed in *Brassica* hexaploid. Of these nonadditive genes, 2,445 showed differential expression either between *Brassica* hexaploid and *B. rapa* or between *Brassica* hexaploid and *B. carinata*. Among the 2,445 differentially expressed nonadditive genes, 1,002 belonged to group 1, 163, to group 2, 43, to group 3, and 1,237, to group 4.

To explore the differences in functional category distribution between additive and nonadditive genes, we assigned them to secondary classification GO terms. The genes belonged to 12, 9, and 23 functional groups in main GO categories of molecular function, cellular component, and biological process, respectively (Figure 9). In the molecular function category, the terms transcription regulator activity (GO:0030528), antioxidant activity (GO:0016209), structural molecular activity (GO:0005198), and catalytic activity (GO:0003824) showed significant differences in gene numbers between additive and nonadditive genes, with transcription regulator activity, antioxidant activity, and catalytic activity being significantly enriched in nonadditive genes. In the cellular component category, the numbers of additive and nonadditive genes differed significantly for the functional terms cell part (GO:0044464), organelle part (GO:0044422), extracellular region part (GO:0044421), organelle (GO:0043226), macromolecular complex (GO:0032991), membrane-enclosed lumen (GO:0031974), cell (GO:0005632), and extracellular region (GO:0005576). Four of these groups (cell part, extracellular region part, cell, and extracellular region) were significantly enriched in nonadditive genes. In the biological process category, nine functional groups, including biological regulation (GO:0065007), multi-organism process (GO:0051704), response to stimulus (GO:0050896), rhythmic process (0048511), pigmentation (GO:0043473), developmental process (GO:
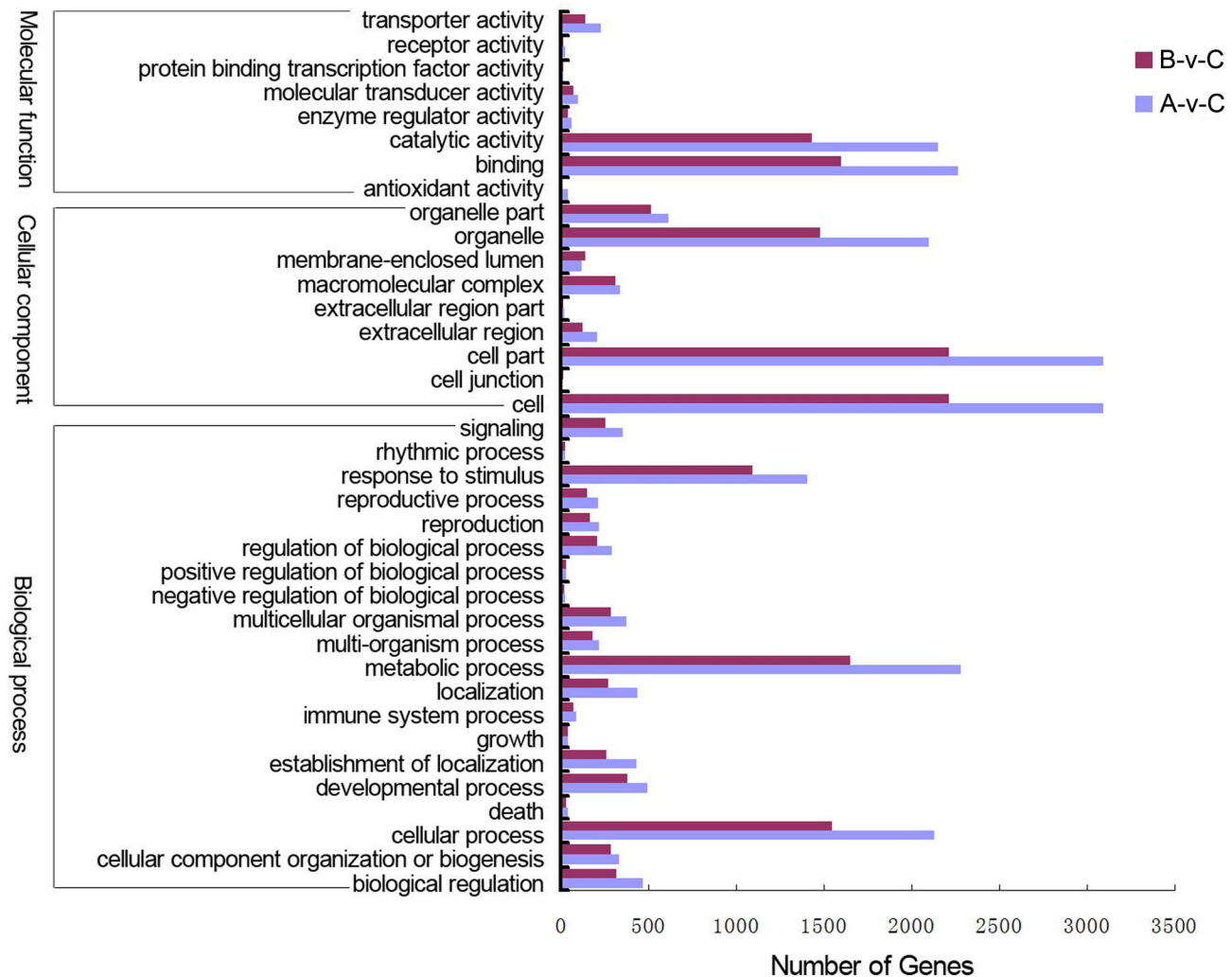
Figure 6. Functional annotation of DEGs between *Brassica* hexaploid and its parents based on GO terms. DEGs between *Brassica* hexaploid (C)-*B. rapa* (A) and *Brassica* hexaploid (C)-*B. carinata* (B) are grouped to the secondary classification of GO terms. There are 8, 9, and 20 functional groups in the three main categories (molecular function, cellular component, and biological process) of GO classification, respectively.

doi: 10.1371/journal.pone.0068883.g006

0032502), cellular process (GO:0009987), metabolic process (0008152), and immune system process (GO:0002376), had significantly more nonadditive genes than additive genes.

Analysis of putative transcription factor and methylase genes

We queried *B. rapa* transcription factor (TF) genes in the Plant Transcription Factor Database (http://planttfdb.cbi.edu.cn/) and identified 589 putative TF genes, which could be classified into 49 families (Table S6). Of the 589 putative TF genes, 143 genes belonging to 36 families showed significant differential expression, accounting for 24.28% of the total TF genes (Table S6). There were 101 putative TF genes differentially expressed between *Brassica* hexaploid and *B. rapa*, with 49 genes upregulated and 52

genes downregulated in *Brassica* hexaploid. Seventy-six putative TF genes showed differential expression between *Brassica* hexaploid and *B. carinata*: 46 genes were upregulated and 30 were downregulated in *Brassica* hexaploid.

The TCP gene family is a family of transcription factors present in various plant species. We detected 15 TCP genes in our RNA-Seq data, and five showed significantly different expression between *Brassica* hexaploid and its parents (Table 3). Three TCP genes (Bra001992, Bra003959, and Bra006616) were expressed at higher levels in *Brassica* hexaploid compared with *B. rapa*, and two genes (Bra003959 and Bra006616) were upregulated in *Brassica* hexaploid compared with *B. carinata*. The other two differentially expressed TCP genes (Bra009679 and Bra011803) showed lower expression levels in *Brassica* hexaploid compared with *B. rapa*. Some
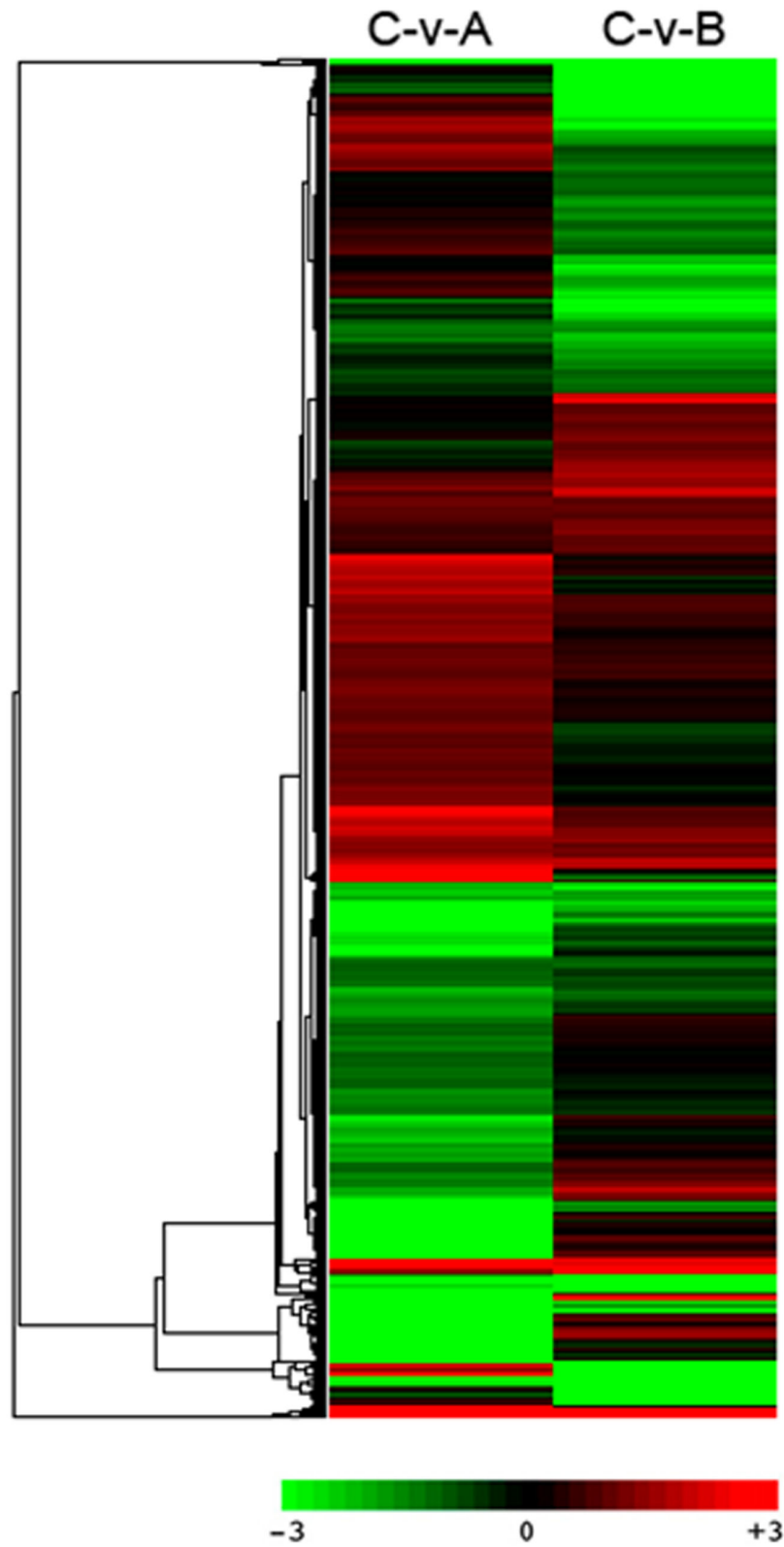
Figure 7. Hierarchical clustering analysis of 7397 DEGs based on log ratio RPKM data. A, *B. rapa*; B, *B. carinata*; C, *Brassica* hexaploid. The column represents individual experiment, and row represents individual gene. Genes up-regulated is represented by red, and genes down-regulated is represented by green.
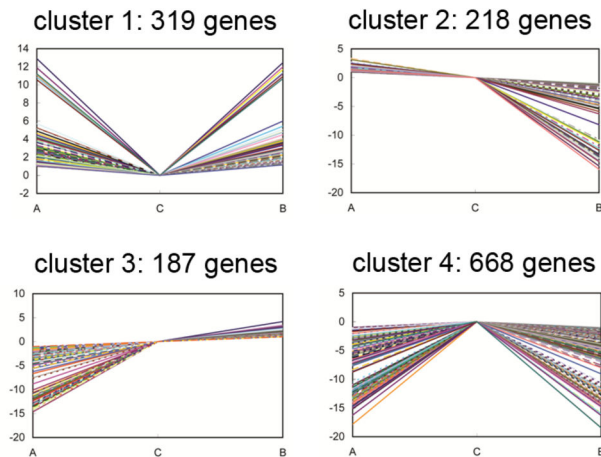
Figure 8. Four clusters of gene expression pattern. A, *B. rapa*; B, *B. carinata*; C, *Brassica* hexaploid. The four clusters of gene expression pattern were grouped by up or down regulation in *Brassica* hexaploid compared to its parents.
doi: 10.1371/journal.pone.0068883.g008

putative target genes of the TCP gene family also showed differential expression. Ten of the 32 genes potentially regulated by TCP genes that we examined showed significant differential expression between *Brassica* hexaploid and its parents.

To explore the underlying mechanism of changes in DNA methylation, we determined the expression levels of putative methyltransferase and methylase genes identified in *Brassica* hexaploid and its parents (Table S7). Among 157 putative methyltransferase genes, 47 (29.94%) exhibited significant differential expression between *Brassica* hexaploid and its parents. Between *Brassica* hexaploid and *B. rapa*, 39 putative methyltransferase genes were differentially expressed in the *Brassica* hexaploid; 32 were upregulated and seven were downregulated. Between *Brassica* hexaploid and *B. carinata*, 16 putative methyltransferase genes were upregulated and four were downregulated. Compared with the number of methyltransferase genes, fewer putative methylase genes were identified, with 11 genes being detected, one of which (Bra035280) was also annotated as a methyltransferase. Among these 11 genes, three (not including Bra035280) were significantly upregulated and one was significantly downregulated in *Brassica* hexaploid compared with *B. rapa*. No putative methylase gene showed significantly different expression between *Brassica* hexaploid and *B. carinata*.

Discussion

RNA-Seq has been used for transcriptomic analysis in a number of plant species [16,30–34]. Here, we used RNA-Seq to trace the transcriptomic changes between *Brassica* hexaploid and its parents. Before bioinformatic analysis could be performed, we required a reference genome for aligning the reads. In previous studies, an *Arabidopsis* 70-mer oligo microarray was used to measure gene expression in

resynthesized *B. napus* [26]. Considering the close relationship between *Arabidopsis* and *Brassica*, we mapped the *Brassica* hexaploid library reads to the *A. thaliana* reference genome (ftp://ftp.arabidopsis.org/Sequences/whole_chromosomes). The total mapped reads accounted for only 21.10% of the library reads, with 17.77% uniquely matched reads. The match scores of the *A. thaliana* genome were too low for transcriptome analysis (data unpublished). We then used the *B. rapa* genome as a reference, and the uniquely matched reads accounted for 70.35%, 46.20%, and 54.18% in the *B. rapa*, *B. carinata*, and *Brassica* hexaploid libraries, respectively. As the match scores fit the technological criterion, the *B. rapa* genome could be applied as a reference. A total of 6.26–7.43% of the multiply mapped reads in our sequencing data remains unanalyzed; however, most of these might have been generated from duplicated genes and chromosome segments. A total of 23.39–46.37% of the reads could not be matched to the *B. rapa* genome. This relatively low percentage may be attributable to the inclusion of species different from *B. rapa*, unavoidable errors in the reference genome, or the specific mapping criteria. Nevertheless, the numerous genes and DEGs detected in *Brassica* hexaploid and its parents provide valuable information for using synthesized *Brassica* hexaploid as a model to explore changes in gene expression.

Parental-biased gene expression and non-additive gene expression in *Brassica* hexaploid

The parental-biased changes in gene expression and DNA methylation between synthesized *B. napus* and its parents, as well as between newly synthesized *Brassica* allohexaploid and its parents, were explored using cDNA-AFLP and MSAP [36,47]. To explore parental-biased gene expression in the allohexaploid, we compared the number and fold-change of DEGs between *Brassica* hexaploid and its maternal parent *B. carinata* and paternal parent *B. rapa*, respectively. Of the 7,397 DEGs identified between *Brassica* hexaploid and its parents, a total of 5,248 DEGs (71%) were differently expressed between *Brassica* hexaploid and *B. rapa*, and 3,644 DEGs (about 50%) showed differential expression between *Brassica* hexaploid and *B. carinata*.

To analyze the magnitudes of the differences in expression, we first divided the absolute difference between the RPKM value in *Brassica* hexaploid and the RPKM value in *B. rapa* by the RPKM value in *Brassica* hexaploid for each DEG. Similarly, we divided the absolute difference between the RPKM values in *Brassica* hexaploid and *B. carinata* by the RPKM value in *Brassica* hexaploid for each DEG. We then compared the normalized values for each gene by subtracting the later from the former, such that a positive final value indicated that the magnitude of the difference in the expression of the gene between *Brassica* hexaploid and *B. rapa* was larger than that between *Brassica* hexaploid and *B. carinata*. A negative value indicated the converse. Of the 7,397 DEGs, 4,425 DEGs (~60%) had a positive value and 2,854 DEGs (~39%) had a negative value. The remaining 52 DEGs (~1%) had a final value of zero. As there were more DEGs with a larger difference in expression between *Brassica* hexaploid and *B. rapa* than between *Brassica* hexaploid and *B. carinata*, it is
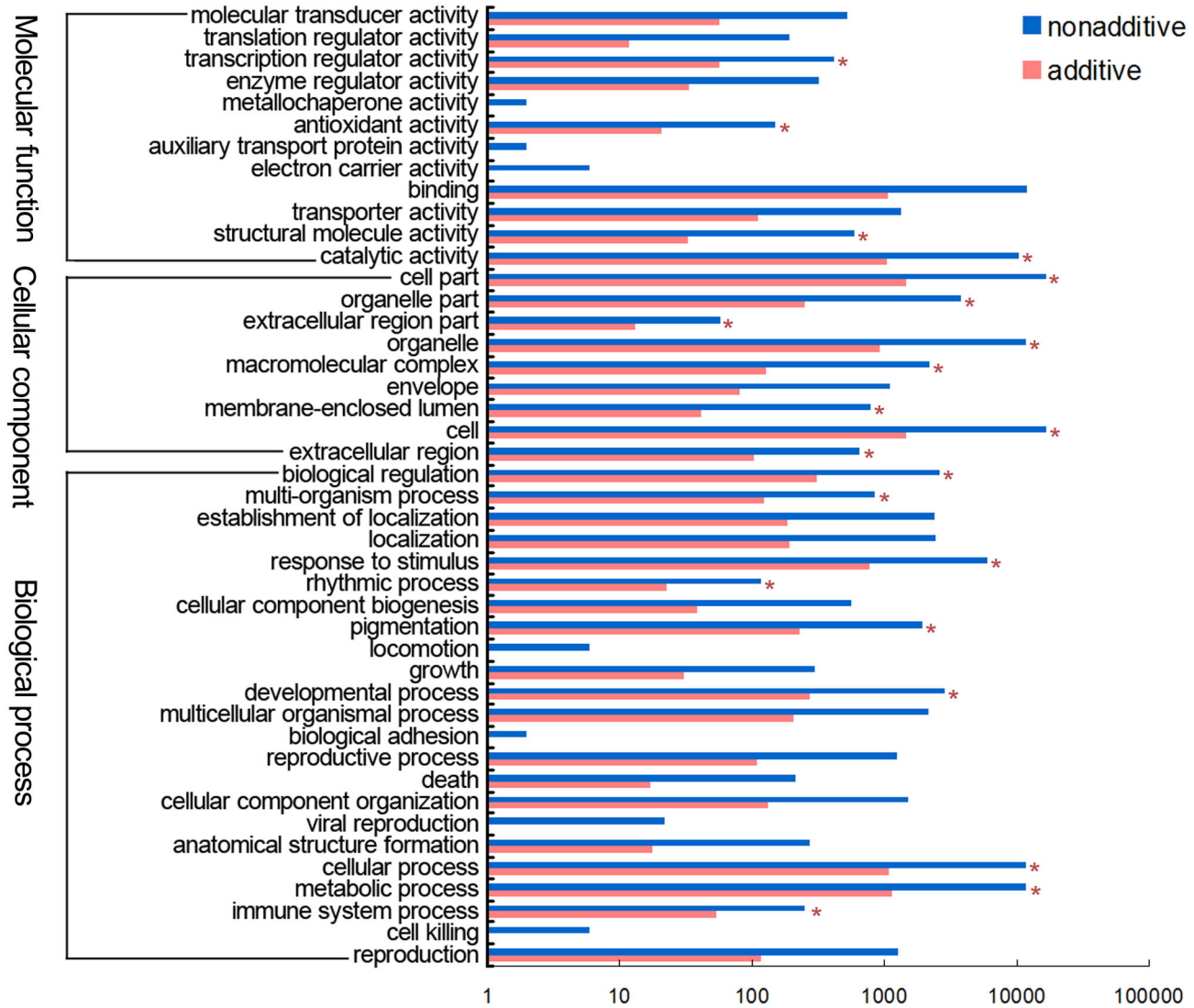
Figure 9. Functional annotation of nonadditive genes and additive genes in *Brassica* hexaploid based on GO terms. Nonadditive genes and additive genes are grouped to the secondary classification of GO terms. There are 12, 9, and 23 functional groups in the three main categories (molecular function, cellular component, and biological process) of GO classification, respectively.

doi: 10.1371/journal.pone.0068883.g009

possible that there were directional gene expression changes away from the paternal parent in *Brassica* hexaploid.

Our results are in accordance with the results of Xu et al. [36,47], which also showed larger gene expression differences between polyploids and their paternal parent. However, for more than 60% of the proteins in allotetraploid *B. napus*, the expression patterns were closer to those of its paternal parent, *B. rapa* [48]. Apart from differences generated by the divergence of experimental materials, it is also conceivable that changes at the transcriptomic level may not agree with changes at the proteomic level. The paternal-biased phenomenon has been shown in other polyploids. For example, Adams et al. [17,21] observed silencing and biased expression of homologous gene pairs in natural and synthetic tetraploid cotton, and noted that gene silencing in the newly synthesized allotetraploid cotton was mainly epigenetic. One possible explanation for paternal-biased changes in allopolyploids is cytoplasmic and maternal effects. Cytoplasm-nuclear interactions could have influenced the parental genomes, resulting in differences in the extent and direction of genomic and transcriptomic changes.

Nonadditive gene expression in polyploidy has been observed in *Arabidopsis* allotetraploids, hexaploid wheat, and citrus allotetraploid hybrids [18,49,50]. The percentage of nonadditive genes in *Brassica* hexaploid was lower than that in hexaploid wheat and higher than that in *Arabidopsis*

Table 3. TCP genes detected in RNA-Seq data.

GeneID	Length	A_RPKM	B_RPKM	C_RPKM
Bra001992	3102	0.364129	1.124772	2.335480
Bra003959	1122	1.901563	4.318980	29.517329
Bra006616	858	5.850964	3.162822	12.263409
Bra009679	486	26.340172	20.340797	10.647648
Bra011803	1242	21.927688	5.306295	10.277295
Bra001887	330	-	-	1.568108
Bra010597	1491	0.168348	-	-
Bra004415	525	0.239054	2.953689	2.628448
Bra004788	1287	0.975161	0.150611	0.938184
Bra003834	1521	4.373220	1.146958	2.494952
Bra008235	684	26.421721	26.638257	30.261735
Bra003890	288	20.917196	2.692164	14.374324
Bra011417	1017	18.140577	10.292168	10.515729
Bra003179	1671	16.373245	15.195986	17.651773
Bra004223	2622	1.292367	2.661361	1.513087

A, *B. rapa*; B, *B. carinata*; C, *Brassica* hexaploid. The five genes marked by underline were significantly differently expressed genes between *Brassica* hexaploid and its parents.

allotetraploids [18,49]. Because RNA-Seq technology is much more precise than microarrays, we could use restrictive criteria to select nonadditive genes. The percentage of nonadditive genes in *Brassica* hexaploid would increase if less restrictive criteria were used. There were no significant relationships between the nonadditive genes in citrus allotetraploids and the functional categories of genes [50]. However, significant numbers of nonadditive genes in *Arabidopsis* allotetraploids belonged to the hormonal regulation, cell defense, and aging categories [18]. Compared with the additive genes in *Brassica* hexaploid, the nonadditive genes were significantly enriched in biological processes related to stimulus response, immune system, cellular metabolic, rhythmic process, developmental process, pigmentation, seed germination, and tropism, and in the molecular functions associated with catalytic activity, antioxidant activity, and transcription regulator activity.

The high frequency of changes in gene expression in the present study was presumably attributable in part to the higher ploidy level in *Brassica* hexaploid. The diversity of the genomes in allohexaploids may decrease the genomic stability compared with that in allotetraploids. Therefore, adjustments in gene expression are necessary for the adaptation of allohexaploids to the environment. The *Brassica* hexaploid can also be considered as a secondary polyploidy derivative of tetraploid *B. carinata*, and the conflict between tetraploid and diploid genomes may lead to more severe changes in the hexaploid transcriptome. We hypothesized that the natural tetraploid *B. carinata* used in our experiment had a lesser impact on the transcriptome when hybridized with the diploid *B. rapa* (for it has previously undergone allopolyploidization), and that the genes in *B. carinata* had entered a stabilized state.

Putative transcription factor and methylase genes may be important in allohexaploid evolution

Transcription factors regulate the expression of many downstream genes and have critical roles in plant development and adaptation. In synthesized *Arabidopsis* allotetraploids, the regulation of genome-wide non-additive gene expression may be partly controlled by transcription factors [18]. Our results provide a comprehensive view of transcription factor gene families. We examined several multi-gene families to identify their expression profiles and functions in plants. As studies on the TCP gene family increase, their importance in the evolution and developmental control of plant forms is being elucidated [51]. One member of the TCP gene family, *teosinte branched 1*, determines the strong apical dominance in domesticated maize [52]. Another gene, *CYCLOIDEA* from snapdragon, controls floral bilateral symmetry [53]. *PROLIFERATING CELL FACTORS 1* and *2* from rice are involved in DNA replication and repair, maintenance of chromatin structure, chromosome segregation, and cell-cycle progression [54]. We detected five TCP genes with significant differential expression between *Brassica* hexaploid and its parents; three were upregulated and two were downregulated. Meanwhile, one-third of the putative target genes of TCP genes were differentially expressed between the hexaploid and its parents. Thus, direct or indirect regulation of the interactions between TCP genes and their target genes may occur. The synergistic actions between members of the TCP gene family and their target genes may contribute to the morphological evolution of polyploids.

Auxin response factors (ARF) are transcription factors that control plant growth and development by regulating the expression of auxin response genes [55]. We detected two ARF genes (Bra007127 and Bra009079), and their expression levels did not differ significantly between *Brassica* hexaploid and its parents. This suggests that the morphological differences that develop between *Brassica* hexaploid and its parents may be determined by genes of multiple functional groups, excluding auxin-related genes.

DNA methylation is an important epigenetic factor for transcriptomic changes in allopolyploids, and changes in DNA methylation and their relationship with gene expression in allopolyploids have been widely studied. Alterations in cytosine methylation affected both repetitive DNA sequences and low-copy DNA in wheat allotetraploids [56]. Significant changes in methylation patterns were observed in *Spartina* allopolyploids and may explain the morphological plasticity and larger ecological amplitude of allopolyploids [25]. In a newly synthesized *Brassica* tetraploid line, *B. napus*, extensively altered DNA methylation was detected using restriction fragment length polymorphism probes [57]. Hegarty et al. [58] observed methylation state changes, which mirrored gene expression changes, in triploid *Senecio* hybrids and their allohexaploid derivatives. In our results, 47 of the 157 putative methyltransferase genes and 4 of the 11 putative methylase genes showed significant differential expression between *Brassica* hexaploid and its parents. Among the differentially expressed genes, 40 were upregulated and 11 were downregulated. Gene methylation leads to gene inactivation; thus, the upregulation of methylation-related genes may be

associated with the lower expression of some DEGs in the hexaploid compared with its parents, and the downregulation of methyltransferase and methylation genes may influence the increased expression of some DEGs. Although the methylation states of genes were not examined in our study, altered expression of DNA methylation-related genes provides indirect evidence for the function of DNA methylation in transcriptome changes during allopolyploid evolution.

Conclusions

This study identified genes expressed in *Brassica* hexaploid and its parents, detected genes differentially expressed between the hexaploid and its parents, and classified these genes according to different functional categories. The *Brassica* hexaploid showed more differences in gene expression compared with its paternal parent than compared with its maternal parent, which supports the paternal-biased changes suggested in many previous allopolyploid studies. The DEGs detected between the hexaploid and its parents revealed that the adaptation and evolution processes of polyploids involve genome-wide changes in gene expression. Many nonadditive genes associated with important biological processes were identified in *Brassica* hexaploid. We speculate that transcriptomic changes are more severe in allohexaploids because of their higher ploidy level and the processes of secondary polyploidization. Changes in the transcriptomes of newly formed allopolyploids provide more variations for selection and facilitate allopolyploid survival and evolution. Further studies on gene regulation are required to determine the molecular mechanisms behind the success of allopolyploids in nature.

Supporting Information

Table S1. A list of KEGG pathways mapped by 21092 genes in our result. Pathways over-represented by genes were for metabolic pathways, biosynthesis of secondary metabolites, plant–pathogen interaction, and plant hormone signal transduction with 4529, 2547, 1867, and 1684 members respectively.
(XLS)

Table S2. A set of 5248 differentially expressed genes between *Brassica* hexaploid and *B. rapa*. FDR: false discovery rate. We screened genes with expression change no less than two-fold and FDR no more than 0.001 as differentially expressed genes. There were 2670 up-regulated genes and 2578 down-regulated genes compared *Brassica* hexaploid (C) to *B. rapa* (A).
(DOC)

Table S3. A set of 3644 differentially expressed genes between *Brassica* hexaploid and *B. carinata*. FDR: false

discovery rate. We screened genes with expression change no less than two-fold and FDR no more than 0.001 as differentially expressed genes. There were 2274 up-regulated genes and 1370 down-regulated genes compared *Brassica* hexaploid (C) to *B. carinata* (B).
(DOC)

Table S4. A list of genes in cluster 1, 2, 3, and 4, respectively. For the 1392 DEGs significantly differential expressed in both comparisons of *Brassica* hexaploid (C)-*B. rapa* (A) and *Brassica* hexaploid (C)-*B. carinata* (B), we classified them into four expression patterns by two-dimensional hierarchical clustering. There were 319, 218, 187, and 668 genes in cluster 1, 2, 3, and 4, respectively.
(XLS)

Table S5. A list of genes in group 1, 2, 3, and 4, respectively. For the 7397 DEGs significantly differential expressed in either comparison of *Brassica* hexaploid (C)-*B. rapa* (A) and *Brassica* hexaploid (C)-*B. carinata* (B), we classified them into four groups. Genes in group 1, 2, 3, and 4 had the same expression profiles as genes in cluster 1, 2, 3, and 4. There were 2016, 1701, 1327, and 2353 genes in group 1, 2, 3, and 4 respectively.
(XLS)

Table S6. A list of putative transcription factors (TFs) in our result. A, *B. rapa*; B, *B. carinata*; C, *Brassica* hexaploid. We queried *B. rapa* transcription factors genes in the Plant Transcription Factor Database (http://planttfdb.cbi.edu.cn/) and identified 589 putative TF genes, which classified into 49 families. Of the 589 putative TFs, 143 genes belonged to 36 families which showed significantly differential expression between *Brassica* hexaploid and its parents are marked by *.
(XLS)

Table S7. A list of the putative methyltransferase and methylase genes in our result.
A, *B. rapa*; B, *B. carinata*; C, *Brassica* hexaploid. We screened 157 putative methyltransferase genes and 11 putative methylase genes with one gene (Bra035280) annotated by both methyltransferase gene and methylase gene. The differentially expressed genes between them are marked by *.
(XLS)

Author Contributions

Conceived and designed the experiments: JW. Performed the experiments: QZ JW. Analyzed the data: QZ. Contributed reagents/materials/analysis tools: JZ JM SM. Wrote the manuscript: QZ JZ JW.

References

1. Jiao Y, Wickett NJ, Ayyampalayam S, Chanderbali AS, Landherr L et al. (2011) Ancestral polyploidy in seed plants and angiosperms. Nature 473: 97-100. doi:10.1038/nature09916. PubMed: 21478875.

2. Blanc G, Hokamp K, Wolfe KH (2003) A recent polyploidy superimposed on older large-scale duplications in the Arabidopsis genome. Genome Res 13: 137-144. doi:10.1101/gr.751803. PubMed: 12566392.

3. Jaillon O, Aury JM, Noel B, Policriti A, Clepet C et al. (2007) The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. Nature 449: 463-467. doi:10.1038/nature06148. PubMed: 17721507.

4. Tang H, Bowers JE, Wang X, Ming R, Alam M et al. (2008) Synteny and collinearity in plant genomes. Science 320: 486-488. doi:10.1126/science.1153917. PubMed: 18436778.

5. Wolfe KH (2001) Yesterday's polyploids and the mystery of diploidization. Nat Rev Genet 2: 333-341. doi:10.1038/35072555. PubMed: 11331899.

6. Soltis PS, Soltis DE (2009) The role of hybridization in plant speciation. Annu Rev Plant Biol 60: 561-588. doi:10.1146/annurev.arplant.043008.092039. PubMed: 19575590.

7. Rieseberg LH, Raymond O, Rosenthal DM, Lai Z, Livingstone K et al. (2003) Major ecological transitions in wild sunflowers facilitated by hybridization. Science 301: 1211-1216. doi:10.1126/science.1086949. PubMed: 12907807.

8. Peloquin SJ, Ortiz R (1992) Plant Breeding in the 1990s: Techniques for introgressing unadapted germ plasm to breeding populations. Stalker HT, Murohy JP. CAB International. pp. 485-507.

9. Osabe K, Kawanabe T, Sasaki T, Ishikawa R, Okazaki K et al. (2012) Multiple mechanisms and challenges for the application of allopolyploidy in plants. Int J Mol Sci 13: 8696-8721. doi:10.3390/ijms13078696. PubMed: 22942729.

10. Comai L (2005) The advantages and disadvantages of being polyploid. Nat Rev Genet 6: 836-846. doi:10.1038/ni1226. PubMed: 16304599.

11. Kashkush K, Feldman M, Levy AA (2012) Gene loss, silencing and activation in a newly synthesized wheat allotetraploid. Genetics 160: 1651-1659.

12. Gaeta RT, Pires JC, Iniguez-Luy F, Leon E, Osborn TC (2007) Genomic changes in resynthesized *Brassica napus* and their effect on gene expression and phenotype. Plant Cell 19: 3403-3417. doi:10.1105/tpc.107.054346. PubMed: 18024568.

13. Bento M, Pereira HS, Rocheta M, Gustafson P, Viegas W et al. (2008) Polyploidization as a retraction force in plant genome evolution: sequence rearrangements in Triticale. PLOS ONE 3: e1402. doi:10.1371/journal.pone.0001402. PubMed: 18167561.

14. Wang Q, Liu H, Gao A, Yang X, Liu W et al. (2012) Intergenomic rearrangements after polyploidization of *Kengyilia thoroldiana* (Poaceae: Triticeae) affected by environmental factors. PLOS ONE 7: e31033. doi:10.1371/journal.pone.0031033. PubMed: 22363542.

15. Kashkush K, Feldman M, Levy AA (2003) Transcriptional activation of retrotransposons alters the expression of adjacent genes in wheat. Nat Genet 33: 102-106. PubMed: 12483211.

16. Zenoni S, Ferrarini A, Giacomelli E, Xumerle L, Fasoli M et al. (2010) Characterization of transcriptional complexity during berry development in *Vitis vinifera* using RNA-Seq. Plant Physiol 152: 1787-1795. doi:10.1104/pp.109.149716. PubMed: 20118272.

17. Adams KL, Percifield R, Wendel JF (2004) Organ-specific silencing of duplicated genes in a newly synthesized cotton allotetraploid. Genetics 168: 2217-2226. doi:10.1534/genetics.104.033522. PubMed: 15371349.

18. Wang J, Tian L, Lee HS, Wei NE, Jiang H et al. (2006) Genomewide nonadditive gene regulation in Arabidopsis allotetraploids. Genetics 172: 507-517. PubMed: 16172500.

19. Fujimoto R, Taylor JM, Sasaki T, Kawanabe T, Dennis ES (2011) Genome wide gene expression in artificially synthesized amphidiploids of Arabidopsis. Plant Mol Biol 77: 419-431. doi:10.1007/s11103-011-9820-y. PubMed: 21882042.

20. Pikaard CS (2001) Genomic change and gene silencing in polyploids. Trends Genet 17: 675-677. doi:10.1016/S0168-9525(01)02545-8. PubMed: 11718903.

21. Adams KL, Cronn R, Percifield R, Wendel JF (2003) Genes duplicated by polyploidy show unequal contributions to the transcriptome and organ-specific reciprocal silencing. Proc Natl Acad Sci U S A 100: 4649-4654. doi:10.1073/pnas.0630618100. PubMed: 12665616.

22. Zhao J, Buchwaldt L, Rimmer SR, Brkic M, Bekkaoui D et al. (2009) Differential expression of duplicated peroxidase genes in the allotetraploid Brassica napus. Plant Physiol Biochem 47: 653-656. doi:10.1016/j.plaphy.2009.03.002. PubMed: 19345111.

23. Rastogi S, Liberles DA (2005) Subfunctionalization of duplicated genes as a transition state to neofunctionalization. BMC Evol Biol 5: 28. doi:10.1186/1471-2148-5-28. PubMed: 15831095.

24. Chaudhary B, Flagel L, Stupar RM, Udall JA, Verma N et al. (2009) Reciprocal silencing, transcriptional bias and functional divergence of homeologs in polyploid cotton (*Gossypium*). Genetics 182: 503-517. doi:10.1534/genetics.109.102608. PubMed: 19363125.

25. Salmon A, Ainouche ML, Wendel JF (2005) Genetic and epigenetic consequences of recent hybridization and polyploidy in *Spartina* (Poaceae). Mol Ecol 14: 1163-1175. doi:10.1111/j.1365-294X.2005.02488.x. PubMed: 15773943.

26. Gaeta RT, Yoo SY, Pires JC, Doerge RW, Chen ZJ et al. (2009) Analysis of gene expression in resynthesized *Brassica napus* Allopolyploids using Arabidopsis 70mer oligo microarrays. PLOS ONE 4: e4760. doi:10.1371/journal.pone.0004760. PubMed: 19274085.

27. Wall PK, Leebens-Mack J, Chanderbali AS, Barakat A, Wolcott E et al. (2009) Comparison of next generation sequencing technologies for transcriptome characterization. BMC Genomics 10: 347. doi:10.1186/1471-2164-10-347. PubMed: 19646272.

28. Weber AP, Weber KL, Carr K, Wilkerson C, Ohlrogge JB (2007) Sampling the Arabidopsis transcriptome with massively parallel pyrosequencing. Plant Physiol 144: 32-42. doi:10.1104/pp.107.096677. PubMed: 17351049.

29. Cheung F, Haas BJ, Goldberg SM, May GD, Xiao Y et al. (2006) Sequencing *Medicago truncatula* expressed sequenced tags using 454 Life Sciences technology. BMC Genomics 7: 272. doi:10.1186/1471-2164-7-272. PubMed: 17062153.

30. Xu H, Gao Y, Wang J (2012) Transcriptomic analysis of rice (*Oryza sativa*) developing embryos using the RNA-Seq technique. PLOS ONE 7: e30646. doi:10.1371/journal.pone.0030646. PubMed: 22347394.

31. Lu T, Lu G, Fan D, Zhu C, Li W et al. (2010) Function annotation of the rice transcriptome at single-nucleotide resolution by RNA-seq. Genome Res 20: 1238-1249. doi:10.1101/gr.106120.110. PubMed: 20627892.

32. Severin AJ, Woody JL, Bolon YT, Joseph B, Diers BW et al. (2010) RNA-Seq atlas of *Glycine max*: a guide to the soybean transcriptome. BMC Plant Biol 10: 160. doi:10.1186/1471-2229-10-160. PubMed: 20687943.

33. Mutasa-Göttgens ES, Joshi A, Holmes HF, Hedden P, Göttgens B (2012) A new RNASeq-based reference transcriptome for sugar beet and its application in transcriptome-scale analysis of vernalization and gibberellin responses. BMC Genomics 13: 99. doi:10.1186/1471-2164-13-99. PubMed: 22429863.

34. Wu T, Qin Z, Zhou X, Feng Z, Du Y (2010) Transcriptome profile analysis of floral sex determination in cucumber. J Plant Physiol 167: 905-913. doi:10.1016/j.jplph.2010.02.004. PubMed: 20303197.

35. Soltis DE, Soltis PS, Tate JA (2003) Advances in the study of polyploidy since plant speciation. New Phytol 161: 173-191. doi:10.1046/j.1469-8137.2003.00948.x.

36. Xu Y, Zhao Q, Mei S, Wang J (2012) Genomic and transcriptomic alterations following hybridisation and genome doubling in trigenomic allohexaploid *Brassica carinata × Brassica rapa*. Plant Biol 14: 734-744. doi:10.1111/j.1438-8677.2011.00553.x.

37. Jiang Y, Tian E, Li R, Chen L, Meng J (2007) Genetic diversity of *Brassica carinata* with emphasis on the interspecific crossability with *B. rapa*. Plant Breeding 126: 487-491. doi:10.1111/j.1439-0523.2007.01393.x.

38. Tian E, Jiang Y, Chen L, Zou J, Liu F et al. (2010) Synthesis of a *Brassica* trigenomic allohexaploid (*B. carinata × B. rapa*) de novo and its stability in subsequent generations. Theor Appl Genet 121: 1431-1440. doi:10.1007/s00122-010-1399-1. PubMed: 20607208.

39. Wang X, Wang H, Wang J, Sun R, Wu J et al. (2011) The genome of the mesopolyploid crop species *Brassica rapa*. Nat Genet 43: 1035-1039. doi:10.1038/ng.919. PubMed: 21873998.

40. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods 5: 621-628. doi:10.1038/nmeth.1226. PubMed: 18516045.

41. Conesa A, Götz S, García-Gómez JM, Terol J, Talón M et al. (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21: 3674-3676. doi:10.1093/bioinformatics/bti610. PubMed: 16081474.

42. Ye J, Fang L, Zheng H, Zhang Y, Chen J et al. (2006) WEGO: a web tool for plotting GO annotations. Nucleic Acids Res 34: W293-W297. doi:10.1093/nar/gkl031. PubMed: 16845012.

43. Kanehisa M, Goto S (2000) KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res 28: 27-30. doi:10.1093/nar/28.1.27. PubMed: 10592173.

44. Bindea G, Mlecnik B, Hackl H, Charoentong P, Tosolini M et al. (2009) ClueGO: a cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. Bioinformatics 25: 1091-1093. doi:10.1093/bioinformatics/btp101. PubMed: 19237447.

45. Marioni JC, Mason CE, Mane SM, Stephens M, Gilad Y (2008) RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. Genome Res 18: 1509-1517. doi:10.1101/gr.079558.108. PubMed: 18550803.

46. Audic S, Claverie JM (1997) The significance of digital gene expression profiles. Genome Res 7: 986-995. PubMed: 9331369.

47. Xu Y, Zhong L, Wu X, Fang X, Wang J (2009) Rapid alterations of gene expression and cytosine methylation in newly synthesized *Brassica napus* allopolyploids. Planta 229: 471-483. doi:10.1007/s00425-008-0844-8. PubMed: 18998158.

48. Albertin W, Balliau T, Brabant P, Chèvre AM, Eber F et al. (2006) Numerous and rapid nonstochastic modifications of gene products in newly synthesized *Brassica napus* allotetraploids. Genetics 173: 1101-1113. doi:10.1534/genetics.106.057554. PubMed: 16624896.

49. Pumphrey M, Bai J, Laudencia-Chingcuanco D, Anderson O, Gill BS (2009) Nonadditive expression of homoeologous genes is established upon polyploidization in hexaploid wheat. Genetics 181: 1147-1157. doi:10.1534/genetics.108.096941. PubMed: 19104075.

50. Bassene JB, Froelicher Y, Dubois C, Ferrer RM, Navarro L et al. (2010) Non-additive gene regulation in a citrus allotetraploid somatic hybrid between *C. reticulata* Blanco and *C. limon* (L.) Burm. Heredity 105: 299-308. doi:10.1038/hdy.2009.162. PubMed: 19953121.

51. Martín-Trillo M, Cubas P (2010) TCP genes: a family snapshot ten years later. Trends Plant Sci 15: 31-39. doi:10.1016/j.tplants.2009.11.003. PubMed: 19963426.

52. Doebley J, Stec A, Hubbard L (1997) The evolution of apical dominance in maize. Nature 386: 485-488. doi:10.1038/386485a0. PubMed: 9087405.

53. Luo D, Carpenter R, Vincent C, Copsey L, Coen E (1996) Origin of floral asymmetry in *Antirrhinum*. Nature 383: 794-799. doi:10.1038/383794a0. PubMed: 8893002.

54. Kosugi S, Ohashi Y (1997) PCF1 and PCF2 specifically bind to cis elements in the rice proliferating cell nuclear antigen gene. Plant Cell 9: 1607-1619. doi:10.2307/3870447. PubMed: 9338963.

55. Guilfoyle TJ, Hagen G (2007) Auxin response factors. Curr Opin Plant Biol 10: 453-460. doi:10.1016/j.pbi.2007.08.014. PubMed: 17900969.

56. Shaked H, Kashkush K, Ozkan H, Feldman M, Levy AA (2001) Sequence elimination and cytosine methylation are rapid and reproducible responses of the genome to wide hybridization and allopolyploidy in wheat. Plant Cell 13: 1749-1759. doi:10.1105/tpc.13.8.1749. PubMed: 11487690.

57. Lukens LN, Pires JC, Leon E, Vogelzang R, Oslach L et al. (2006) Patterns of sequence loss and cytosine methylation within a population of newly resynthesized *Brassica napus* allopolyploids. Plant Physiol 140: 336-348. PubMed: 16377753.

58. Hegarty MJ, Batstone T, Barker GL, Edwards KJ, Abbott RJ et al. (2011) Nonadditive changes to cytosine methylation as a consequence of hybridization and genome duplication in *Senecio* (Asteraceae). Mol Ecol 20: 105-113. doi:10.1111/j.1365-294X.2010.04926.x. PubMed: 21073590.