

Changes in genome-wide gene expression during allopolyploidization and genome stabilization in hexaploid wheat

Yeonju Jung¹, Kanako Kawaura¹, Kohei Mishina¹, Shun Sakuma¹,
Masahiro Kishii² and Yasunari Ogihara^{1*}

¹Kihara Institute for Biological Research and Department of Life and Environmental System Science,
Yokohama City University, Maioka-cho 641-12, Totsuka-ku, Yokohama 244-0813, Japan

²International Maize and Wheat Improvement Center (CIMMYT), Km 45 Carretera México-Veracruz El Batán,
Texcoco, Mexico CP 56130

(Received 7 October 2014, accepted 11 November 2014)

Allopolyploidization is an important evolutionary event in plants, but its genome-wide effects are not fully understood. Common wheat, *Triticum aestivum* (AABBDD), evolved through amphidiploidization between *T. turgidum* (AABB) and *Aegilops tauschii* (DD). Here, global gene expression patterns in the seedlings of a synthetic triploid wheat line (ABD), its chromosome-doubled hexaploid (AABBDD) and stable synthetic hexaploid (AABBDD), and the parental lines *T. turgidum* (AABB) and *Ae. tauschii* (DD) were compared using an oligo-DNA microarray to identify metabolic pathways affected by the genome conflict that occurs during allopolyploidization and genome stabilization. Characteristic gene expression patterns of non-additively expressed genes were detected in the newly synthesized triploid and hexaploid, and in the stable synthetic hexaploid. Hierarchical clustering of all differentially expressed and non-additively expressed genes revealed that the gene expression patterns of the triploid (ABD) were similar to those of the maternal parent (AABB), and that expression patterns in successive generations arising from self-pollination became closer to that of the pollen parent (DD). The non-additive gene expression profiles markedly differed between the triploid (ABD) and chromosome-doubled hexaploid (AABBDD), as supported by Gene Ontology (GOSlim) analysis. Four hundred and nineteen non-additively expressed genes were commonly detected in all three generations. GOSlim analysis indicated that these non-additively expressed genes were predominantly involved in “biological pathways”. Notably, four of 11 genes related to sugar metabolism displayed elevated expression throughout allopolyploidization. These may be useful candidates for promoting heterosis and adaptation in plants.

Key words: allopolyploidization, DNA microarray, gene expression, heterosis, hexaploid wheat

INTRODUCTION

Allopolyploidization results from the interspecific or intergeneric hybridization and multiplication of more than two sets of genomes, and is a common and critical event in plant evolution (Soltis et al., 2007; Leitch and Leitch, 2008). The widespread occurrence of allopolyploidization in plants suggests that it provides an evolutionary advantage through speciation and environmental adaptation (Wendel, 2000; Doyle et al., 2008; Soltis and Soltis, 2009). Specifically, the compatible merging of

diverged genomes leads to functional differentiation of duplicated genes (Blanc and Wolfe, 2004; Buggs et al., 2011) and thereby generates novel gene networks and the formation of new traits, which are stably maintained in the descendants of allopolyploids (Osborn et al., 2003; Chen, 2010; Miller et al., 2012). The “genome shock” resulting from the introduction of distinct genomes, and subsequent “genome tuning”, which is required for the stable maintenance of homoeologous genomes, introduce a broad range of genetic and epigenetic responses, such as chromosome deletions, rearrangements, transposable element activation and epigenetic modifications (Pikaard, 1999; Comai et al., 2000; Wendel, 2000; Ozkan et al.,

Edited by Koji Murai

* Corresponding author. E-mail: yogihara@yokohama-cu.ac.jp

2001; Adams et al., 2003; He et al., 2003; Kashkush et al., 2003; Osborn et al., 2003; Wang et al., 2004; Buggs et al., 2011; Ng et al., 2012; Shi et al., 2012).

Allopolyploidization proceeds *via* two main mechanisms: hybridization of unreduced gametes, and chromosome doubling in somatic cells (Stebbins, 1947; Hegarty et al., 2008). The former process was first reported by Karpechenko (1927), who observed that fertile plants among intergeneric hybrids of *Raphanus sativus* ($2n = 2x = 18$) and *Brassica oleracea* ($2n = 2x = 18$) were allotetraploid. Chromosome doubling was detected in interspecific hybrids between *Primula verticillata* ($2n = 2x = 18$) and *P. floribunda* ($2n = 2x = 18$), which spontaneously produced branches that contained tetraploid cells, as is found in *P. kewensis* ($2n = 4x = 36$; Newton and Pellew, 1929). Although overall gene expression has been examined in two systems, one involving crossings between autotetraploids for the production of allotetraploids using *Arabidopsis* (Wang et al., 2006) and the other involving genomic doubling after interspecific hybridization with cotton (Rapp et al., 2009), systematic comparisons of overall gene expression patterns during the process of allohexaploidization have been limited (Hegarty et al., 2006; Chagué et al., 2010; Qi et al., 2012).

Common wheat, *Triticum aestivum* ($2n = 6x = 42$, genomic constitution AABBDD), originated about 10,000 years ago through two successive allopolyploidization events. The first involved hybridization between the wild relatives *Aegilops speltoides* ($2n = 2x = 14$, SS $\div$ BB) and *T. urartu* ($2n = 2x = 14$, AA), and occurred about 0.5 million years ago. The resultant tetraploid, *T. turgidum* ssp. *dicoccoides* ($2n = 4x = 28$, AABB), was domesticated (*T. turgidum* ssp. *dicoccum*) and hybridized with wild goat grass, *Ae. tauschii* ssp. *strangulata* ($2n = 2x = 14$, DD), generating *T. aestivum* (Kihara, 1944; Feuillet et al., 2008). Common wheat is therefore a useful model for studying the process of allopolyploidization leading to the generation of allohexaploids. The use of wheat to study allopolyploidization is also advantageous because hybridization between allotetraploids (AABB) and goat grass (DD) to obtain a synthetic allohexaploid (AABBDD) is relatively simple. Furthermore, a number of synthetic allohexaploid wheat lines that were generated by self-pollination and have been stably maintained for over 50 years are readily available.

Several studies have examined genome-wide events in newly formed hexaploid wheat lines. For example, using a 17-K feature microarray, Pumphrey et al. (2009) found that 16.2% of transcripts (134/825) displayed non-additive expression in a newly generated allohexaploid wheat, while Chagué et al. (2010) reported that 7% of transcripts (3,853/55,049) on the Affymetrix GeneChip Wheat Genome array were expressed non-additively in newly formed wheat allohexaploids. Notably, the observed expression patterns were stable in both lines and between

generations. Using the Affymetrix GeneChip Wheat Genome array, Akhunova et al. (2010) also documented that 19% of transcripts showed non-additive expression. More recently, Qi et al. (2012) used two nascent allohexaploid wheat lines that shared the same diploid parental genotype to identify genes with altered expression and found that 2.8% to 5.5% of genes represented in the Affymetrix GeneChip Wheat Genome array showed non-additive expression, which could be classified based on the transgenerational heritability and inter-line conservation of expressed genes. Although global gene expression patterns during allopolyploidization in wheat have been studied, only early generations of allohexaploids were analyzed. Furthermore, the allohexaploid lines were generated using cultivated wheat tetraploids with different genotypes as maternal parents (Feuillet et al., 2008), and no studies have examined genome-wide gene expression profiles in hybridized triploids (ABD). In addition, the direct relationship between gene expression changes during the different stages of interspecific hybridization, including genome doubling during allopolyploidization, and those that occur during genome stabilization after successive self-pollinations using the same combination of parents, remain unclear.

Here, in an attempt to better understand the genome-wide changes in gene expression that occur during allopolyploidization and genome stabilization, we examined the global gene expression patterns in the seedlings of newly formed synthetic triploids and of their chromosome-doubled and stable allohexaploids, which were stably maintained through self-pollination for 13 generations using the same combination of parents. We used a 38K oligo-DNA microarray (Agilent Technologies) to investigate the molecular basis of heterosis for biomass production in the vegetative stage (Tang et al., 2011). Gene expression levels in synthetic polyploids were compared with mid-parent values (MPVs) averaged from the parental species to determine the non-additively expressed gene expression patterns. Using this approach, we identified several changes in the non-additively expressed gene expression patterns, especially those of genes involved in sugar metabolism, that occur during allopolyploidization and genome stabilization of *T. aestivum*.

MATERIALS AND METHODS

Plant materials and growth conditions *Triticum turgidum* ssp. *dicoccum* (accession no. KU124, AABB) and *Ae. tauschii* ssp. *strangulata* (accession no. KU2074, DD) were used to generate allopolyploids. Global gene expression profiles among the newly generated synthetic triploid (ABD-S₀) and its doubled hexaploid (AABBDD-S₁), stable synthetic hexaploid (accession no. KU-221-13, AABBDD-S₁₄), *T. turgidum* ssp. *dicoccum*, and *Ae. tauschii* were compared (Supplementary Fig. S1). The

first generation of the synthetic hexaploid KU-221-13 was produced in 1959 by the late Prof. Tanaka (Kyoto University) by crossing KU124 (AABB) and KU2074 (DD), and was self-pollinated for 13 generations at Kyoto University. Plants of the S₁₄ generation of KU-221-13 (AABBDD-S₁₄), KU124 (AABB) and KU2074 (DD), which were provided by the National Bioresource Project (NBRP; Laboratory of Crop Evolution, Graduate School of Agriculture, Kyoto University), were analyzed in the present study. To duplicate the original allopolyploidization event, we generated new synthetic lines by crossing KU124 (AABB) as the female parent with KU2074 (DD) as the pollen parent to generate the triploid (ABD-S₀). Triploid embryos were rescued 2 weeks after pollination and were cultured in petri dishes containing half strength Murashige and Skoog culture medium (Inagaki et al., 1998) in a growth chamber maintained at 20°C with a 12-h light/12-h dark cycle (5000 lux). The S₁ generation of allohexaploid plants (AABBDD-S₁) was obtained by the self-pollination of colchicine-treated (0.2% (w/v) in a 2% (v/v) DMSO solution) triploid plants (ABD-S₀). All plants were grown in a growth chamber maintained at 22°C and 40% relative humidity, with a 16-h light (6000 lux; 64,560 fc) and 8-h dark cycle. For RNA extraction, triploid seeds without embryo rescue and colchicine treatment were used.

RNA extraction and microarray analysis Total RNA was extracted in triplicate from two-leaf seedlings using a Plant RNeasy Mini Kit (QIAGEN) according to the manufacturer's method. Extracted RNAs (500 ng) were labeled using an Agilent Low RNA Input Fluorescent Linear Amplification kit (Agilent Technologies) with cyamine 3-cytidine 5C-triphosphate (CTP) for hybridization with a 60-mer oligonucleotide microarray comprising 37,886 probes (Kawaura et al., 2008). Each hybridization treatment was independently performed in three biological replicates. The hybridized microarrays were scanned using a DNA microarray scanner (Agilent Technologies) and signal intensities were detected from the obtained digital images using Feature Extraction software (Agilent Technologies).

Microarray data analysis and statistics Fluorescence intensities were analyzed with GeneSpring GX software (v12.6; Agilent Technologies). Raw data sets (raw signal intensity > 1.0 threshold) were transformed to log scale (base 2). Normalization of the signals to the 75th percentile for all arrays and to the median for all probes was performed with GeneSpring. Weakly hybridized probe spots were stringently removed from the analysis using a cut-off value for the raw signal intensity in all microarray slides of < 20%. Low-quality hybridizing probes, which were identified based on their detection *P*-values determined using GeneSpring GX (v12.6), were also removed. Mid-parent values (MPVs) between *T. turgidum* (AABB)

and *Ae. tauschii* (DD) were estimated *in silico* by averaging the expression values across the treatments, with the ratio of *T. turgidum* and *Ae. tauschii* data set as 1:1 (Chagué et al., 2010; Pumphrey et al., 2009; Rapp et al., 2009). Similarities in gene expression patterns between lines were estimated using Pearson's correlation coefficient, as described by Eisen et al. (1998). For the comparison of the two parental lines, Welch's *t*-test (*P* < 0.01) was applied to detect expression differences between the two lines using GeneSpring GX (v12.6). To control for multiple testing errors, the false discovery rate (FDR) of Benjamini and Hochberg (1995) was employed and controlled the FDR below the level of 0.05. One-way analysis of variance (ANOVA) was used for determining non-additive gene expression profiles based on the Benjamini and Hochberg multiple testing correction procedure to control the false discovery rate (FDR = 0.05). The hybridization data obtained from the oligo-DNA microarray analyses were deposited in the NCBI Gene expression Omnibus (GSE54398).

Gene ontology was analyzed using the original sequences of microarray probes, which were obtained from the NBRP KOMUGI database (<http://www.shigen.nig.ac.jp/wheat/komugi/>). ORF regions were predicted using getORF (version 1.5, JemBoss Package), and the estimated peptide sequences were assigned to gene ontology categories using Interproscan (5-44.0) with the Pfam database. GO annotations were determined by BLASTx searches (v.2.2.26.) of barley (<http://mips.helmholtz-muenchen.de/plant/barley/>), rice (TIGR rice mcu v7, <http://rice.plantbiology.msu.edu/>), *Brachypodium* (Gramene, <http://www.gramene.org/>) and sorghum (Gramene, <http://www.gramene.org/>) using threshold *E*-values of < 1E-50, < 1E-20, < 1E-20 and < 1E-10, respectively. For KEGG pathway analysis, the original sequences of microarray probes were assigned to rice pseudomolecules (TIGR BLASTx, *E*-value < 1e-20); the obtained IDs were converted to RAP IDs using the RAP-DB ID converter (<http://rapdb.dna.affrc.go.jp/tools/converter>), and the genes were mapped to KEGG pathway *Oryza sativa japonica* (RAPDB) databases.

Real-time quantitative RT-PCR For cDNA synthesis, 1 µg total RNA was incubated for 5 min at 65°C with 10 mM Oligo-(dT)20 and dNTPs. Reverse transcription (RT) was then performed with the ReverTra Ace enzyme (Toyobo) for 1 h at 42°C, after which the reaction mixture was heated for 5 min at 96°C to inactivate the enzyme. Quantitative real-time PCR was performed for six randomly selected genes in each hierarchical cluster. cDNA was amplified using 1× SYBR Premix Ex Taq II (Takara) in a 20-µl reaction mixture using the following conditions: denaturation for 10 s at 95°C, followed by 40 cycles of 10 s at 95°C and 30 s at 60°C. The threshold cycle (Ct) of the primary amplification curve was used for calculations.

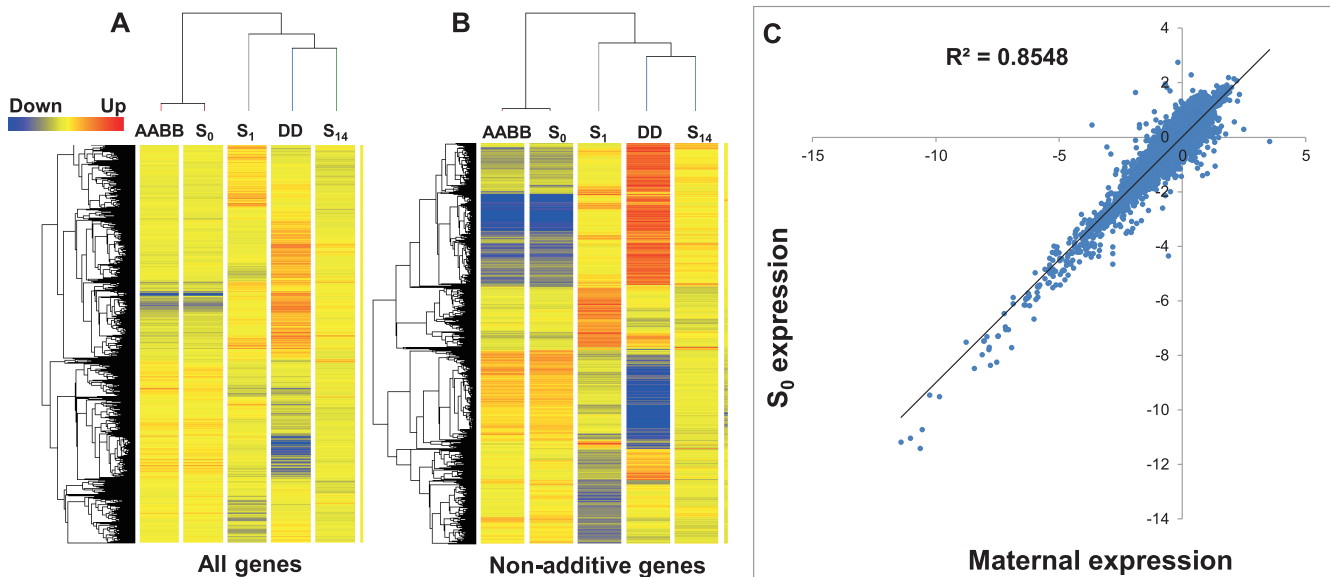


Fig. 1. Hierarchical clustering of expressed genes in the three generations of synthetic polyploids (ABD-S₀, AABBDD-S₁ and AABBDD-S₁₄) and their parental lines. Hierarchical clustering of all 20,802 genes (A) and 3,799 non-additively expressed genes (B). Expression patterns of the genes in three allopolyploids and their parental lines are also shown. Degree of expression (up- and down-regulation) is indicated by the color gradient. (C) Scatter plot of the relative expression of all detected genes in ABD-S₀ and the maternal parent AABB.

The quantity of each product was normalized to that of the actin gene (Ning et al., 2013). Each qPCR was independently performed for at least two biological replicates.

RESULTS

Genome-wide analysis of non-additive gene expression in synthetic polyploids Global gene expression levels of the synthetic polyploid lines (ABD-S₀, AABBDD-S₁ and AABBDD-S₁₄) as well as their parental lines AABB and DD were analyzed using a 38K- oligonucleotide microarray for common wheat (Kawaura et al., 2008). Two-leaf stages were chosen for RNA extraction to facilitate comparison with data from previous work (Tang et al., 2011). To determine transcriptome divergence in allopolyploids, we introduced MPV as the control: if allopolyploidization has no effect, a given gene in the allopolyploid might be expressed equally to the average of its parents (MPV), and would be defined as an additively expressed gene; on the other hand, a non-additively expressed gene would display a value different from the MPV. Non-additive gene expression for three synthetic polyploid lines, a synthetic triploid (ABD-S₀), a synthetic hexaploid (AABBDD-S₁) and a stable synthetic hexaploid (AABBDD-S₁₄), was evaluated (Supplementary Fig. S2). The chromosome numbers of ABD-S₀, AABBDD-S₁ and AABBDD-S₁₄ were confirmed by observing chromosomes in root tip cells (data not shown). Among the three synthetic polyploids, 3,799 transcripts exhibited non-additive gene expression profiles with levels more than two-fold higher and lower than MPVs. Since each generation of

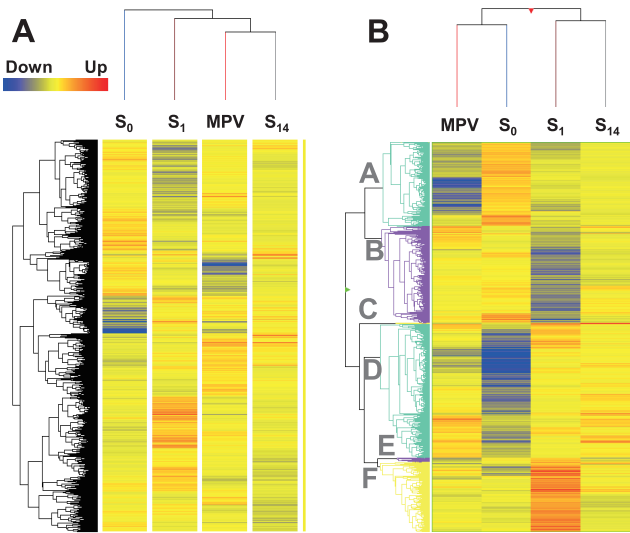


Fig. 2. Hierarchical clustering of expressed genes in the three synthetic lines and MPVs between two parental lines. Hierarchical clustering of all 20,802 genes (A) and 3,799 non-additively expressed genes (B). Expression patterns of the genes in three allopolyploids and the mid-parent values (MPVs) are also shown. Degree of expression (up- and down-regulation) is indicated by the color gradient. In (B), A to F represent clusters showing similar gene expression patterns. MPVs were obtained using the data of AABB and DD.

synthetic wheat showed characteristic expression profiles, gene regulation was substantially altered in the interspecific hybrid and subsequent chromosome doubling, and stabilized during successive self-pollination.

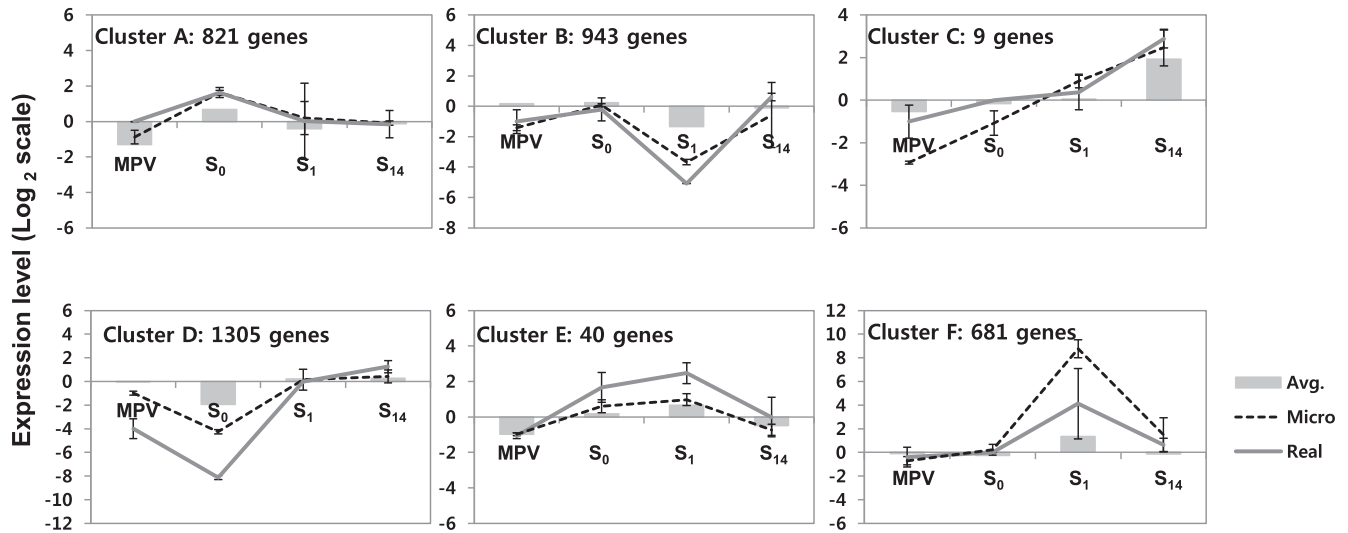


Fig. 3. Expression patterns of six non-additively expressed gene groups classified from hierarchical clustering. Averages of the $\log_2$ -transformed ratios and standard deviations are shown. Data for all genes clustered into each group are shown in the histogram (Avg.; bars). Expression patterns of selected genes determined by DNA microarray (Micro; broken line) and real-time RT-PCR analyses (Real; solid line) are also displayed.

Gene expression profiles in synthetic and progenitor lines To evaluate the overall gene expression patterns in the progenitor and synthetic lines, 20,802 of the 37,886 probes selected by quality control and 3,799 non-additively expressed genes were clustered using a hierarchical clustering method (Figs. 1 and 2). The overall expression patterns of the detected genes among the three synthetic and progenitor lines are shown in Fig. 1A and C. Although the gene expression pattern in ABD- S_0 was similar to that of the maternal parent AABB, the pattern in the synthetic line AABBDD- S_1 was closer to that in DD than to that in AABB. Notably, the gene expression pattern of AABBDD- S_{14} was relatively close to that of the pollen parent DD (Fig. 1A). The same trends were also observed in the expression patterns of non-additively expressed genes (Fig. 1B).

The observed profiles of differential gene expression were supported by the MPVs between the tetraploid (AABB) and diploid (DD) lines (Fig. 2). The MPVs of all expressed genes were closer to the gene expression levels in AABBDD- S_{14} than to those in ABD- S_0 and AABBDD- S_1 (Fig. 2A). In contrast, the overall expression levels of non-additively expressed genes in S_0 (ABD) were closer to the MPVs than were those of the hexaploids S_1 and S_{14} (Fig. 2B). These data also suggested that the gene expression profile in the triploid line (ABD) initially resembled that in AABB (the maternal parent), but became closer to the MPV during successive pollinations, and that the cluster of non-additively expressed genes also changed after allopolyploidization.

Characterization of expression patterns of non-additively expressed genes during hexaploidization Gene expression patterns of non-additively expressed genes were classified into six major groups based on comparisons with the MPV (Fig. 2B). These groups were designated as clusters A to F. The relative expression patterns and MPV of non-additively expressed genes among three generations are shown in Fig. 3. The 821 genes in cluster A were specifically up-regulated in ABD- S_0 . The 943 genes in cluster B showed down-regulation in AABBDD- S_1 . The smallest cluster, C, includes nine genes, which were up-regulated in all three generations; their expression level gradually increased from ABD- S_0 to AABBDD- S_{14} . The 1,305 genes belonging to the largest cluster, D, showed down-regulation in ABD- S_0 compared to MPV and up-regulation in the two hexaploids S_1 and S_{14} . The expression level of the 40 genes in cluster E gradually increased from ABD- S_0 to AABBDD- S_1 compared to MPV but was similar to MPV in AABBDD- S_{14} . The 681 genes in cluster F were up-regulated only in AABBDD- S_1 . Relatively fewer expressed genes in the parental lines tended to be affected. To verify the microarray data, we performed real-time RT-PCR for one randomly selected gene from each of the six clusters (Supplementary Table S1). These six genes showed non-additive expression in at least one of the allopolyploid generations. Actin was used as a control for quantification. The expression levels of these six genes determined by real-time RT-PCR were consistent with the microarray results of the selected genes (Fig. 3), suggesting the reliability of the microarray data.

To gain insight into heritability of the expression over generations, the numbers of non-additively expressed

genes identified in the three synthetic generations, ABD-S₀, AABBDD-S₁ and AABBDD-S₁₄, were compared (Fig. 4). Among the 2,276 non-additively expressed genes detected in ABD-S₀, 1,469 (64.5%) were not found in the other two generations, indicating that the genes were specifically expressed in the first generation. Similarly, 1,444 (67.7%) non-additively expressed genes found in AABBDD-S₁ were generation-specific, whereas only 26 (3.9%) of 668 non-additively expressed genes in AABBDD-S₁₄ were generation-specific. A total of 419 genes were non-additively expressed in all three synthetic generations. Taken together, these results suggest that for most of the non-additively expressed genes, their differential expression was not sustained throughout successive pollinations.

To identify the functional categories of the genes non-

additively expressed during the hybridization and genome doubling, genes that were expressed specifically in the ABD-S₀ and AABBDD-S₁ generations were analyzed by the Gene Ontology (GOSlim) classification method. GO annotations of the microarray probes were compared with those of four reference organisms, barley, *Brachypodium*, sorghum and rice, whose GO data were obtained from the Gramene database (<http://www.gramene.org/>). The proportion of annotated probes corresponding to biological process and cellular component categories on the wheat microarray was similar to those of *Brachypodium*, sorghum and rice. Probes corresponding to genes related to molecular function were similarly represented among the wheat microarray and four reference species (Supplementary Fig. S3). The GO annotation analysis suggests that the wheat microarray is representative of the functional categories and relative proportions of genes found in the four reference species. Therefore, we categorized the non-additively expressed genes that were specifically expressed in the ABD-S₀ and AABBDD-S₁ generations within the biological process, cellular component and molecular function categories. The classifications of the non-additively expressed genes in each of the three categories markedly differed between the two generations (Fig. 5). For example, up-regulated genes grouped as “response to abiotic stimulus” were the most abundant in the “biological process” category in ABD-S₀, but were the least abundant in AABBDD-S₁. The “ribosome” and “structural molecule activity” groups contained the most genes in the cellular component and

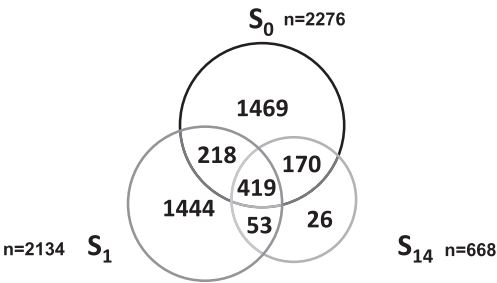


Fig. 4. Venn diagram of non-additively expressed genes detected in the three generations of synthetic polyploids. Overlapping regions indicate genes that were commonly detected in the different lines. The total gene number in each generation is given as n.

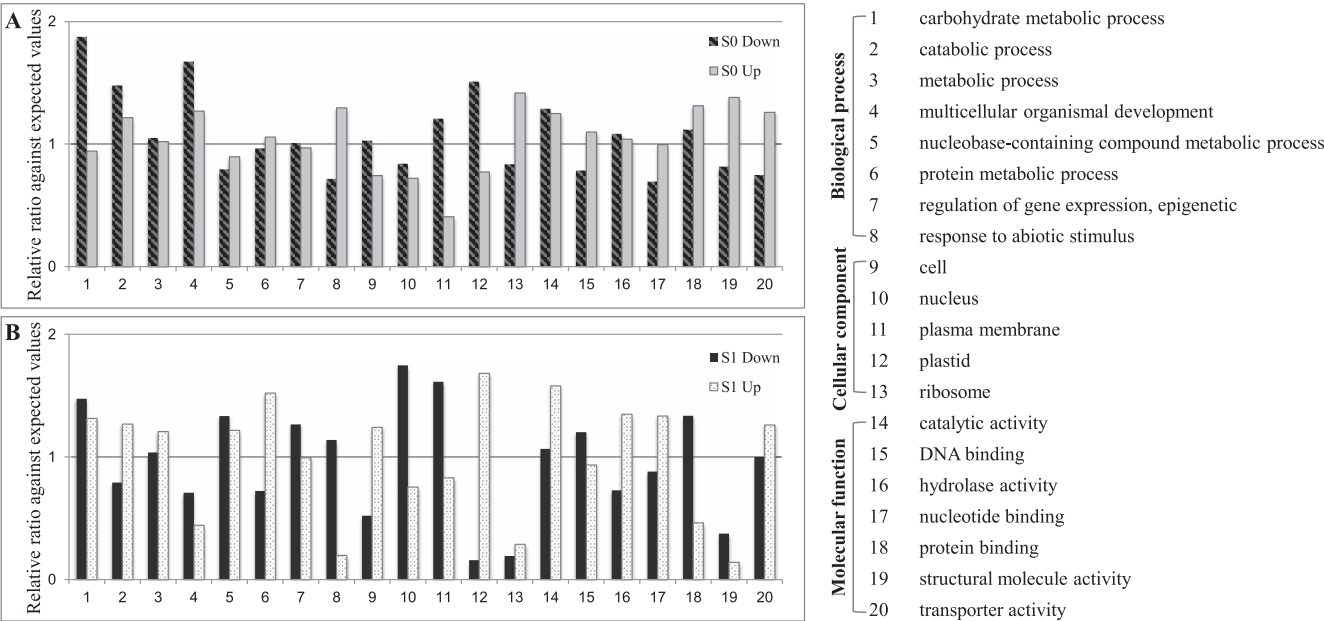


Fig. 5. Gene ontology classification (GOSlim) of non-additively expressed genes that were specifically expressed in S₀ (A) and S₁ (B) generations of the synthetic polyploids. Relative ratio of number of genes belonging to each category is shown against the expected value. Dark- and light-colored bars are down- and up-regulated genes, respectively.

molecular function categories, respectively, in ABD-S₀, but were only minor groups in AABBDD-S₁. Similar trends were found among the down-regulated genes, although fewer differences in the GO classification patterns between generations were observed (Fig. 5).

Molecular characterization of non-additively expressed genes commonly detected in the three synthetic generations To characterize non-additively expressed genes that were commonly detected after hybridization between the tetraploid and diploid lines, we

performed GOSlim analysis for the genes identified among the three synthetic generations. Approximately 11% (419/3,799) of the non-additively expressed genes were commonly found in all three generations and could be divided into five groups based on their expression patterns over successive generations (Fig. 6A). In the first group, 69.5% (291/419) of the non-additively expressed genes showed up-regulation compared to MPV in all three generations. In the second most abundant group, which comprised 28.2% (118/419) of the genes, down-regulation was observed in the S₀ generation, whereas up-regulated

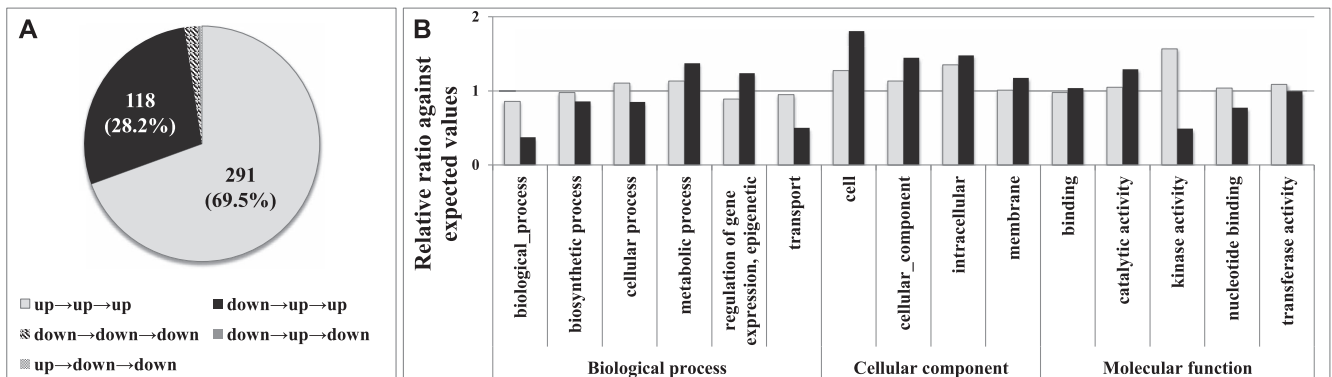


Fig. 6. Characterization of non-additively expressed genes commonly detected throughout the three generations of synthetic polyploids. The five main expression patterns in each of the three generations (S₀, S₁ and S₁₄) are indicated (A). Gene ontology classification of non-additively expressed genes with the two major expression patterns (up-up-up [gray bars] and down-up-up [black bars]) are shown (B).

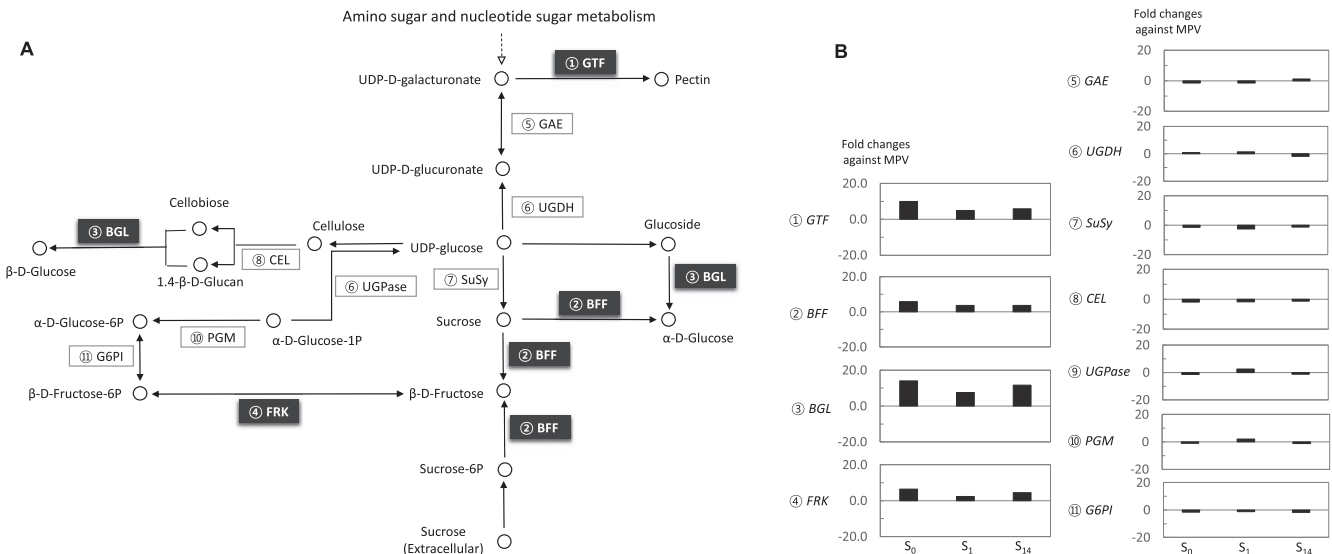


Fig. 7. Expression patterns of genes encoding enzymes related to sugar metabolism. (A) Schematic diagram of rice sugar metabolism. The pathway shown is part of the KEGG starch and sucrose metabolism pathway (KEGG: dosa00500). Non-additively up-regulated genes (GTF, BFF, BGL and FRK) in all three generations are black-colored. The nomenclature used in the KEGG pathway database is adopted. The four identified enzymes were glycosyltransferase (GTF; Os08g0327100, EC:2.4.1.43), β-fructofuranosidase (BFF; Os02g0106100, EC:3.2.1.26), β-glucosidase (BGL; Os03g0749300, EC:3.2.1.21) and fructokinase (FRK; Os08g0113100, EC:2.7.1.4). Abbreviations, annotations and connections are presented in accordance with KEGG standards: solid lines represent a direct relationship, the dashed line represents an indirect relationship and arrowheads indicate directionality. Circles represent metabolites and rectangles indicate enzymes. (B) Expression patterns of 11 genes related to sucrose metabolism. Expression patterns of the four genes that were non-additively up-regulated in all three generations are shown on the left.

expression was detected in the two allohexaploid generations. Genes that were consistently down-regulated in all three generations represented a minor fraction (1.9%) of the total genes, while genes with the other two expression patterns only represented 0.2% of all genes (Fig. 6A).

We further categorized the non-additively expressed genes found in the two dominant expression profile groups using GOSlim. In the first group, genes categorized as 'metabolic processes' in the GO biological process category, 'intercellular' and 'cell' in the GO cellular component category, and 'kinase activity' in the GO molecular function category were most overrepresented (Fig. 6B). In the second group, genes categorized as 'metabolic processes' and 'regulation of gene expression epigenetic' in the GO biological process category, and 'catalytic activity' in the GO molecular function category, were overrepresented. The genes categorized into the GO cellular component category tended to be enriched compared to the other two categories (Fig. 6B). These data suggest that genes related to metabolism and cellular processes are more highly regulated during allopolyploidization and genome stabilization.

Since ploidy levels had effects on growth vigor and starch content in *A. thaliana* (Miller et al., 2012), genes for starch and sucrose synthesis were surveyed. On the basis of KEGG pathway classification for starch and sucrose metabolism (KEGG: dosa00500), expression patterns of eleven of the 20,802 genes in the microarray related to sugar metabolism were analyzed during allopolyploidization (Fig. 7). Four genes displayed elevated expression levels in all generations of allopolyploidization (ABD-S₀, AABBDD-S₁ and AABBDD-S₁₄). These four genes encode glycosyltransferase (Os08g0327100, EC:2.4.1.43), β -fructofuranosidase (Os02g0106100, EC:3.2.1.26), fructokinase (Os08g0113100, EC:2.7.1.4) and β -glucosidase (Os03g0749300, EC:3.2.1.21). Stimulation of sugar metabolism due to increased expression of these genes may enhance heterosis in allohexaploids.

DISCUSSION

Gene regulation during allohexaploidization in common wheat Drastic changes of gene expression patterns are frequently observed in allopolyploids produced by interspecific hybridization rather than in autopolyploids (Birchler et al., 2003; Wang et al., 2006). We systematically examined the genome-wide gene expression profiles in seedlings during the allohexaploidization of wheat. We are interested in the molecular basis of heterosis in the vegetative growth stage to increase biomass production. Therefore, we have focused on the transcriptomics in the seedlings of allopolyploid wheats. The two parental lines of the synthesized allohexaploids (AABB and DD) showed markedly diverged gene expres-

sion profiles, as in the reported case of wheat (Chagué et al., 2010; Qi et al., 2012) and other plant species such as cotton (Rapp et al., 2009) and *Arabidopsis* (Wang et al., 2006). In addition, the gene expression profiles of the triploid (ABD-S₀), its nascent hexaploid (AABBDD-S₁) and the synthetic line (AABBDD-S₁₄) showed a number of notable differences from those of the parental lines and MPVs. Relatively higher numbers of non-additively expressed genes, which deviated from the MPV, were detected in the S₀ (2,276; 10.9%) and S₁ generations (2,134; 10.3%), whereas markedly fewer such genes were found in the S₁₄ generation (668; 3.2%) as shown in Supplementary Fig. S2. The observed gene expression profiles strongly suggest that genome shock (McClintock, 1984) occurred during the early stages of allopolyploid hybridization and chromosome doubling (Ozkan et al., 2001), and that subsequent genome stabilization events reduced non-additively expressed gene expression during successive self-pollinations (Chagué et al., 2010; Qi et al., 2012). Although numerous non-additively expressed genes were detected in the S₀ generation (ABD-S₀), the overall gene expression in this line was closely correlated with that of the maternal parent (AABB; Fig. 1). Our analyses also indicated that the effects of the pollen parent (DD) gradually increased during successive generations, as the gene expression patterns of the stabilized synthetic hexaploid (AABBDD-S₁₄) were relatively close to the MPV (Fig. 2A). These findings suggest that gene regulation system(s) that specifically suppressed expression of the pollen parent genome were active in the triploid line (ABD-S₀), and that the suppression of these genes was relieved and/or weakened in the chromosome-doubled hexaploid line (AABBDD-S₁₄). Transcriptome shock in the triploid was also reported in flower buds of *Senecio* (Hegarty et al., 2006), in which amelioration occurred after genome duplication rather than drastic changes of overall gene expression after amphidiploidization. This discrepancy between transcriptomes in two newly-generated hexaploids might be due to the tissues examined (young seedling vs. flower bud) and/or genome constitutions (Hegarty et al., 2006, 2008). Preferential suppression of either genome usually takes place in interspecific hybrids and their amphidiploids (Wang et al., 2006; Flagel et al., 2008). Although sudden reunification of divergent genomes may induce genome instability (Adams et al., 2003; Wang et al., 2004; Doyle et al., 2008; Buggs et al., 2011), hybrid- or allopolyploidy-induced incompatibilities may be "settled" by gene expression modulation through chromatin modifications (Osborn et al., 2003; Chen et al., 2004; Ha et al., 2011), transcription factors (Barbash et al., 2003; Shi et al., 2012) and/or RNA interference in the selfing progeny (Wang et al., 2006; Ng et al., 2012). Our comparison of the genome-wide gene expression patterns between a synthetic triploid line and its genome-doubled allohexaploid and stable hexaploid

lines has revealed changes arising from hybridization, genome doubling and genome stabilization during allopolyploidization. Since gene expression profiles in hybrid wheats are still unclear, further investigations are needed to clarify the mechanism(s) for gene regulation of allohexaploid wheat.

Shuffling of gene expression patterns in newly synthesized hexaploid wheat Approximately two-thirds (> 60%) of the non-additively expressed genes detected in the triploid (ABD-S₀) and its nascent hexaploid (AABBDD-S₁) were generation-specific (Fig. 4), indicating that the regulation and expression of non-additively expressed genes markedly differ between the two lines. The results of the GO classification analysis of non-additively expressed genes in each generation support this speculation (Fig. 5). Among up-regulated genes, those related to stress responses, carbohydrate metabolic processes and plant structure were the most abundant annotated during the hybridization process (ABD-S₀), but their abundance diminished after chromosome doubling (AABBDD-S₁). Similar changes in the expression profile were observed for down-regulated genes. These patterns of genome regulation in allohexaploids, which have not yet been reported in wheat, support the idea that interspecific hybridization leads to “genome shock” when the genomes of different lines are introduced, and that chromosome doubling induces “genome tuning”, which affects genome regulation among homoeologous genomes.

Gene expression for sugar metabolism is raised throughout genome hybridization and amphidiploidization The comparison of non-additively expressed genes among the three generations of the synthetic triploid (ABD-S₀), nascent hexaploid (AABBDD-S₁) and stable hexaploid lines (AABBDD-S₁₄) revealed that 419 genes were commonly expressed. The 419 genes represented 18.4%, 18.7% and 67.7% of the differentially expressed genes in each generation, respectively, and approximately 70% of the genes were up-regulated in all three generations (Fig. 6). Nearly all of the remaining genes (approximately 28%) were also up-regulated in a later generation (hexaploid stage; Fig. 6A). The proportion of common non-additively expressed genes between two hexaploid generations was slightly lower than that reported in synthetic wheat allohexaploids (Chagué et al., 2010), but were equivalent to those of Qi et al. (2012). It should be emphasized that our data were used to identify non-additively expressed genes that were expressed throughout three generations: an interspecific hybrid (ABD-S₀), its chromosome-doubled derivative (AABBDD-S₁) and a stable allohexaploid (AABBDD-S₁₄). These genes may be good candidates for promoting hybrid vigor (Birchler et al., 2003; Miller et al., 2012) and improving fitness and selective adaptations (Wang et al., 2004, 2006; Doyle et

al., 2008). GO analysis of these genes (Fig. 6) indicates that although those in the ‘metabolism’ and ‘cellular process’ categories tended to be overrepresented, the candidate genes are dispersed among diverse GO categories and play various roles in allohexaploid wheat. Furthermore, gene expression profiles related to sugar metabolism were surveyed, based on the KEGG pathway classification. Four of eleven genes that regulate the sugar metabolism were up-regulated in all three generations examined, ABD-S₀, AABBDD-S₁ and AABBDD-S₁₄ (Fig. 7). Increased sugar and/or starch content due to ploidy and hybridity effects has been correlated with heterosis (Korn et al., 2008; Miller et al., 2012). Based on these studies, we suggest that a sugar-related mechanism is a candidate pathway to give rise to heterosis in allopolyploid wheat (Fig. 7). Since several more underlying mechanisms, such as cis-trans interactions (Tirosh et al., 2009; Ng et al., 2011), epigenetic regulation (Madlung et al., 2002; Ni et al., 2009; Shen et al., 2012), transposable elements (Kashkush et al., 2003; Kraitshtein et al., 2010; Yaakov and Kashkush, 2011; Zhao et al., 2011) and small RNAs (Ha et al., 2009; Kenan-Eichler et al., 2011), seem to control allopolyploidy-regulated gene expression, further studies, especially on the mechanisms of up-regulation of genes for sucrose metabolism, are needed to determine the key roles and functions of the identified candidate genes in relation to increasing hybrid vigor and adaptation in wheat (Ng et al., 2012).

Hybridization data obtained in oligo-DNA microarray analyses were deposited in the NCBI Gene expression Omnibus (GSE54398). This work was partially supported by Grants-in-Aid for scientific research in the priority area “Comparative Genomics” from the Ministry of Education, Culture, Sports, Science and Technology of Japan. This is contribution no. 1017 from the Kihara Institute for Biological Research, Yokohama City University.

REFERENCES

- Adams, K. L., Cronn, R., Percifield, R., and Wendel, J. F. (2003) Genes duplicated by polyploidy show unequal contributions to the transcriptome and organ-specific reciprocal silencing. *Proc. Natl. Acad. Sci. USA* **100**, 4649–4654.
- Akhunova, A. R., Matniyazov, R. T., Liang, H., and Akhunov, E. D. (2010) Homoeolog-specific transcriptional bias in allopolyploid wheat. *BMC Genomics* **11**, 505.
- Barbash, D. A., Siino, D. F., Tarone, A. M., and Roote, J. (2003) A rapidly evolving MYB-related protein causes species isolation in *Drosophila*. *Proc. Natl. Acad. Sci. USA* **100**, 5302–5307.
- Benjamini, Y., and Hochberg, Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B* **57**, 289–300.
- Birchler, J. A., Auger, D. L., and Riddle, N. C. (2003) In search of the molecular basis of heterosis. *Plant Cell* **15**, 2236–2239.
- Blanc, G., and Wolfe, K. H. (2004) Functional divergence of duplicated genes formed by polyploidy during *Arabidopsis* evolution. *Plant Cell* **16**, 1679–1691.

- Buggs, R. J. A., Zhang, L., Miles, N., Tate, J. A., Gao, L., Wei, W., Schnable, P. S., Barbazuk, B., Soltis, P. S., and Soltis, D. E. (2011) Transcriptomic shock generates evolutionary novelty in a newly formed, natural allopolyploid plant. *Curr. Biol.* **21**, 551–556.
- Chagué, V., Just, J., Mestiri, I., Balzergue, S., Tanguy, A. M., Huneau, C., Huteau, V., Belcram, H., Coriton, O., Jahier, J., et al. (2010) Genome-wide gene expression changes in genetically stable synthetic and natural wheat allohexaploids. *New Phytol.* **187**, 1181–1194.
- Chen, Z. J. (2010) Molecular mechanisms of polyploidy and hybrid vigor. *Trends Plant Sci.* **15**, 57–71.
- Chen, Z. J., Wang, J. L., Tian, L., Lee, H. S., Wang, J. Y. J., Chen, M., Lee, J. S. J., Josefsson, C., Madlung, A., Watson, B., et al. (2004) The development of an *Arabidopsis* model system for genome-wide analysis of polyploidy effects. *Bio. J. Linn. Soc.* **82**, 689–700.
- Comai, L., Tyagi, A. P., Winter, K., Holmes-Davis, R., Reynolds, S. H., Stevens, Y., and Byers, B. (2000) Phenotypic instability and rapid gene silencing in newly formed *Arabidopsis* allotetraploids. *Plant Cell* **12**, 1551–1567.
- Doyle, J. J., Flagel, L. E., Paterson, A. H., Rapp, R. A., Soltis, D. E., Soltis, P. S., and Wendel, J. F. (2008) Evolutionary genetics of genome merger and doubling in plants. *Ann. Rev. Genet.* **42**, 443–461.
- Eisen, M. B., Spellman, P. T., Brown, P. O., and Botstein, D. (1998) Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. USA* **95**, 14863–14868.
- Feuillet, C., Langridge, P., and Waugh, R. (2008) Cereal breeding takes a walk on the wild side. *Trends Genet.* **24**, 24–32.
- Flagel, L., Udall, J., Nettleton, D., and Wendel, J. (2008) Duplicate gene expression in allopolyploid *Gossypium* reveals two temporally distinct phases of expression evolution. *BMC Biol.* **6**, 16.
- Ha, M., Lu, J., Tian, L., Ramachandran, V., Kasschau, K. D., Chapman, E. J., Carrington, J. C., Chen, X., Wang, X. J., and Chen, Z. J. (2009) Small RNAs serve as a genetic buffer against genomic shock in *Arabidopsis* interspecific hybrids and allopolyploids. *Proc. Natl. Acad. Sci. USA* **106**, 17835–17840.
- Ha, M., Ng, D. W. K., Li, W. H., and Chen, Z. J. (2011) Coordinated histone modifications are associated with gene expression variation within and between species. *Genome Res.* **21**, 590–598.
- He, P., Friebe, B. R., Gill, B. S., and Zhou, J. M. (2003) Allopolyploidy alters gene expression in the highly stable hexaploid wheat. *Plant Mol. Biol.* **52**, 401–414.
- Hegarty, M. J., Barker, G. L., Wilson, I. D., Abbott, R. J., Edwards, K. J., and Hiscock, S. J. (2006) Transcriptome shock after interspecific hybridization in *Senecia* is ameliorated by genome duplication. *Curr. Biol.* **16**, 1652–1659.
- Hegarty, M. J., Barker, G. L., Brennan, A. C., Edwards, K. J., Abbott, R. J., and Hiscock, S. J. (2008) Changes to gene expression associated with hybrid speciation in plants: further insights from transcriptomic studies in *Senecia*. *Phil. Trans. R. Soc. B* **363**, 3055–3069.
- Inagaki, M. N., Pfeiffer, W. H., Mergoum, M., and Mujeeb-Kazi, A. (1998) Variation of the crossability of durum wheat with maize. *Euphytica* **104**, 17–23.
- Karpechenko, G. D. (1927) The production of polyploidy gametes in hybrids. *Hereditas* **9**, 349–368.
- Kashkush, K., Feldman, M., and Levy, A. A. (2003) Transcriptional activation of retrotransposons alters the expression of adjacent genes in wheat. *Nat. Genet.* **33**, 102–106.
- Kawaura, K., Mochida, K., and Ogihara, Y. (2008) Genome-wide analysis for identification of salt-responsive genes in common wheat. *Funct. Integr. Genomics* **8**, 277–286.
- Kenan-Eichler, M., Leshkowitz, D., Tal, L., Noor, E., Melamed-Bessudo, C., Feldman, M., and Levy, A. A. (2011) Wheat hybridization and polyploidization results in deregulation of small RNAs. *Genetics* **188**, 263–272.
- Kihara, H. (1944) Discovery of the DD-analyser, one of the ancestors of *Triticum vulgare*. *Agric. Hort.* **19**, 13–14.
- Korn, M., Peterek, S., Mock, H., Heyer, A. G., and Hincha, D. (2008) Heterosis in the freezing tolerance, and sugar and flavonoid contents of crosses between *Arabidopsis thaliana* accessions of widely varying freezing tolerance. *Plant Cell Environ.* **31**, 813–827.
- Kraitshtein, Z., Yaakov, B., Khasdan, V., and Kashkush, K. (2010) Genetic and epigenetic dynamics of a retrotransposon after allopolyploidization of wheat. *Genetics* **186**, 801–812.
- Leitch, A. R., and Leitch, I. J. (2008) Genomic plasticity and the diversity of polyploidy plants. *Science* **320**, 481–483.
- Madlung, A., Masuelli, R. W., Watson, B., Reynolds, S. H., Davison, J., and Comai, L. (2002) Remodeling of DNA methylation and phenotypic and transcriptional changes in synthetic *Arabidopsis* allotetraploids. *Plant Physiol.* **129**, 733–746.
- McClintock, B. (1984) The significance of responses of the genome to challenge. *Science* **226**, 792–801.
- Miller, M., Zhang, C., and Chen, Z. J. (2012) Ploidy and hybridity effects on growth vigor and gene expression in *Arabidopsis thaliana* hybrids and their parents. *G3* **2**, 505–513.
- Newton, W. C. F., and Pellew, C. (1929) *Primula kewensis* and its derivatives. *J. Genet.* **20**, 405–467.
- Ng, D. W., Zhang, C., Miller, M., Palmer, G., Whiteley, M., Tholl, D., and Chen, Z. J. (2011) *cis*- and *trans*-regulation of miR163 and target genes confers natural variation of secondary metabolites in two *Arabidopsis* species and their allopolyploids. *Plant Cell* **23**, 1729–1740.
- Ng, D. W., Lu, J., and Chen, Z. J. (2012) Big roles for small RNAs in polyploidy, hybrid vigor, and hybrid incompatibility. *Curr. Opin. Plant Biol.* **15**, 154–161.
- Ni, Z., Kim, E.-D., Ha, M., Lackey, E., Liu, J., Zhang, Y., Sun, Q., and Chen, Z. J. (2009) Altered circadian rhythms regulate growth vigour in hybrids and allopolyploids. *Nature* **457**, 327–331.
- Ning, S., Wang, N., Sakuma, S., Pourkheirandish, M., Wu, J., Matsumoto, T., Koba, T., and Komatsuda, T. (2013) Structure, transcription and post-transcriptional regulation of the bread wheat orthologs of the barley cleistogamy gene *Cly1*. *Theor. Appl. Genet.* **126**, 1273–1283.
- Osborn, T. C., Pires, J. C., Birchler, J. A., Auger, D. L., Chen, Z. J., Lee, H. S., Comai, L., Madlung, A., Doerge, R. W., Colot, V., et al. (2003) Understanding mechanisms of novel gene expression in polyploids. *Trends Genet.* **19**, 141–147.
- Ozkan, H., Levy, A. A., and Feldman, M. (2001) Allopolyploidy-induced rapid genome evolution in the wheat (*Aegilops-Triticum*) group. *Plant Cell* **13**, 1735–1747.
- Pikaard, C. S. (1999) Nucleolar dominance and silencing of transcription. *Trends Plant Sci.* **4**, 478–483.
- Pumphrey, M., Bai, J., Laudencia-Chingcuanco, D., Anderson, O., and Gill, B. S. (2009) Nonadditive expression of homoeologous genes is established upon polyploidization in hexaploid wheat. *Genetics* **181**, 1147–1157.
- Qi, B., Huang, W., Zhu, B., Zhong, X., Guo, J., Zhao, N., Xu, C., Zhang, H., Pang, J., Han, F., et al. (2012) Global transgenerational gene expression dynamics in two newly synthesized allohexaploid wheat (*Triticum aestivum*) lines. *BMC*

- Biol. **10**, 3.
- Rapp, R. A., Udall, J. A., and Wendel, J. F. (2009) Genomic expression dominance in allopolyploids. *BMC Biol.* **7**, 18.
- Shen, H., He, H., Li, J., Chen, W., Wang, X., Guo, L., Peng, Z., He, G., Zhong, S., Qi, Y., et al. (2012) Genome-wide analysis of DNA methylation and gene expression changes in two *Arabidopsis* ecotypes and their reciprocal hybrids. *Plant Cell* **24**, 875–892.
- Shi, X., Ng, D. W., Zhang, C., Comai, L., Ye, W., and Chen, Z. J. (2012) *Cis*- and *trans*-regulatory divergence between progenitor species determines gene-expression novelty in *Arabidopsis* allopolyploids. *Nat. Commun.* **3**, 950.
- Soltis, P. S., and Soltis, D. E. (2009) The role of hybridization in plant speciation. *Annu. Rev. Plant Biol.* **60**, 561–588.
- Soltis, D. E., Soltis, P. S., Schamske, D. W., Hancock, J. F., Thompson, J. N., Husband, B. C., and Judd, W. S. (2007) Autopolyploidy in angiosperms: have we grossly underestimated the number of species? *Taxon* **56**, 13–30.
- Stebbins, G. L. (1947) Types of polyploids; their classification and significance. *Adv. Genet.* **1**, 403–429.
- Tang, J. W., Ohyama, K., Kawaura, K., Hashinokuchi, H., Kamiya, Y., Suzuki, M., Muranaka, T., and Ogihara, Y. (2011) A new insight into application for barley chromosome addition lines of common wheat: achievement of stigmasterol accumulation. *Plant Physiol.* **157**, 1555–1567.
- Tirosh, I., Reikhav, S., Levy, A. A., and Barkai, N. (2009) A yeast hybrid provides insight into the evolution of gene expression regulation. *Science* **324**, 659–662.
- Wang, J., Tian, L., Madlung, A., Lee, H. S., Chen, M., Lee, J. J., Watson, B., Kagochi, T., Comai, L., et al. (2004) Stochastic and epigenetic changes of gene expression in *Arabidopsis* polyploids. *Genetics* **167**, 1961–1973.
- Wang, J., Tian, L., Lee, H. S., Wei, N. E., Jiang, H., Watson, B., Madlung, A., Osborn, T. C., Doerge, R. W., Comai, L., and Chen, Z. J. (2006) Genomewide nonadditive gene regulation in *Arabidopsis* allotetraploids. *Genetics* **172**, 507–517.
- Wendel, J. F. (2000) Genome evolution in polyploids. *Plant Mol. Biol.* **42**, 225–249.
- Yaakov, B., and Kashkush, K. (2011) Massive alterations of the methylation patterns around DNA transposons in the first four generations of a newly formed wheat allohexaploid. *Genome* **54**, 42–49.
- Zhao, N., Zhu, B., Li, M., Wang, L., Xu, L., Zhang, H., Zheng, S., Qi, B., Han, F., and Liu, B. (2011) Extensive and heritable epigenetic remodeling and genetic stability accompany allohexaploidization of wheat. *Genetics* **188**, 499–509.