

Comparison of genome-wide gene expression patterns in the seedlings of nascent allohexaploid wheats produced by two combinations of hybrids

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

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

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

(Received 20 February 2015, accepted 10 March 2015)

Allopolyploidization in plants is an important event that enhances heterosis and environmental adaptation. Common wheat, *Triticum aestivum* (AABBDD), which is an allohexaploid that evolved from an allopolyploidization event between *T. turgidum* (AABB) and *Aegilops tauschii* (DD), shows more growth vigor and wider adaptation than tetraploid wheats. To better understand the molecular basis for the heterosis of hexaploid wheat, we systematically analyzed the genome-wide gene expression patterns of two combinations of newly hybridized triploids (ABD), their chromosome-doubled hexaploids (AABBDD), stable synthetic hexaploids (AABBDD) and natural hexaploids, in addition to their parents, *T. turgidum* (AABB) and *Ae. tauschii* (DD), using a microarray to reconstruct the events of allopolyploidization and genome stabilization. Overall comparisons of gene expression profiles showed that the newly generated hexaploids exhibited gene expression patterns similar to those of their maternal tetraploids, irrespective of hybrid combination. With successive generations, the gene expression profiles of nascent hexaploids became less similar to the maternal profiles, and belonged to a separate cluster from the natural hexaploids. Triploids revealed characteristic expression patterns, suggesting endosperm effects. In the newly hybridized triploids (ABD) of two independent synthetic lines, approximately one-fifth of expressed genes displayed non-additive expression; the number of these genes decreased with polyploidization and genome stabilization. Approximately 20% of the non-additively expressed genes were transmitted across generations throughout allopolyploidization and successive self-pollinations, and 43 genes overlapped between the two combinations, indicating that shared gene expression patterns can be seen during allohexaploidization. Furthermore, four of these 43 genes were involved in starch and sucrose metabolism, suggesting that these metabolic events play key roles in the hybrid vigor of hexaploid wheat.

Key words: allopolyploidization, DNA microarray, hexaploid wheat, maternal effects, non-additive gene expression

INTRODUCTION

The widespread modes of polyploid speciation in plants suggest that it confers advantages in development, adaptation and diversification (Leitch and Leitch, 2008; Soltis and Soltis, 2009; Soltis et al., 2014). Taxonomists encounter more allopolyploids than autopolyploids in nature (Mallet, 2007; Rieseberg and Willis, 2007; Yakimowski and Rieseberg, 2014), suggesting that heterozygosity and

intergenomic interactions among allopolyploids resulted in more phenotypic variation that provided enhanced growth vigor and wider environmental adaptation; this is known as “heterosis” (Ni et al., 2009). The merging of distinct genomes into one cell may lead to rapid and important changes at the genetic and epigenetic levels (Ozkan et al., 2001; Shaked et al., 2001), such as chromosomal rearrangement, gene deletion, mutation, gene silencing, transposable element activities, small-interfering RNA activities and epigenetic remodeling (Ni et al., 2009; Ha et al., 2011; Ng et al., 2011, 2012; Yaakov and

Edited by Koji Murai

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

Kashkush, 2011; Zhao et al., 2011; Shen et al., 2012; Soltis et al., 2014).

Common wheat, *Triticum aestivum* ($2n = 6x = 42$, AABBDD genomes), resulted from two rounds of allopolyploidization. About 0.5 million years ago, wild emmer wheat *T. turgidum* ssp. *dicoccoides* ($2n = 4x = 28$, AABB) arose from the intergenic hybridization between wild relatives of *Aegilops speltoides* ($2n = 2x = 14$, SS $\div$ BB genome) and *T. urartu* ($2n = 2x = 14$, AA). Wild emmer wheat was grown so that early cultivated tetraploid wheat *T. turgidum* ssp. *dicoccum* ($2n = 4x = 28$, AABB) could be domesticated. Free-threshing tetraploid wheat was produced through mutations of cultivated emmer wheat. Primitive hexaploid wheat was spontaneously generated about 10,000 years ago from an intergenic hybridization between free-threshing tetraploid wheat ($2n = 4x = 28$, AABB) and certain type(s) of wild goat grass, *Ae. tauschii* ($2n = 2x = 14$, DD; Kihara, 1944; Feuillet et al., 2008; Dvorak et al., 2012; Nishijima et al., 2014), following chromosome doubling of unreduced gametes (Kihara et al., 1950). Consequently, common wheat is widely cultivated all over the world, because it shows more heterosis, such as growth vigor, environmental adaptability and disease resistance, than tetraploids (Dubcovsky and Dvorak, 2007).

Recent work suggests that the genes in multiplied genomes make unequal contributions to genetic traits (genome dominance or genome subfractionation; Wang et al., 2006; Schnable et al., 2011; Cheng et al., 2012). Genomic asymmetry for agronomically important traits has been reported among the three genomes (A, B and D) of common wheat (Feldman et al., 2012). Although Chinese Spring wheat (a natural hexaploid wheat) is estimated to have lost 10,000 to 16,000 genes after allohexaploidization (Brenchley et al., 2012), the structural changes found in the hexaploid genomes had mainly occurred at the tetraploid stage (Zhang et al., 2013). Furthermore, based on the wheat syntenome, Pont et al. (2013) reported overdominance of the D genome against the tetraploid genome (AABB) after the second allohexaploidization. Allohexaploid wheats can be artificially reproduced by crossing the unreduced gametes of tetraploid wheats (maternal progenitor) and goat grass (pollen progenitor; Kihara et al., 1950; Tanaka, 1961; Zhang et al., 2010) and/or by colchicine treatment-induced chromosome doubling of hybridized triploids (McFadden and Sears, 1944). Triploid wheat plants can be obtained through seeds or embryo rescue (Inagaki et al., 1998; Jung et al., 2014).

Gene expression changes on allohexaploidization have been reported using nascent allohexaploids. Most reports used microarray approaches that compare the gene expression levels of the allohexaploids to the averaged level of their parents (Pumphrey et al., 2009; Akhunova et al., 2010; Chagué et al., 2010; Qi et al., 2012;

Chelaifa et al., 2013); however, only early generations of allohexaploids were analyzed, and not the F_1 hybrid, i.e., the triploids (ABD), except in one instance (Hatano et al., 2012). The overall gene expression profile of natural allohexaploid wheat reflects the effects of intergenic hybridization, amphidiploidization and genome stabilization during the course of its evolution. Therefore, the genome-wide gene expression profiles of hybridized triploids (ABD) are indispensable for understanding allopolyploidization in wheat.

In this study, we aimed to profile the overall gene expression patterns of wheat during the course of allohexaploidization and genome stabilization. We used the Agilent 60-mer oligo-DNA microarray, which contains about 38,000 different transcripts, to characterize the changes in gene expression of newly produced triploids (ABD), their chromosome-doubled hexaploids (AABBDD), their stable synthetic hexaploids (AABBDD), naturally cultivated hexaploids (AABBDD), and their parental lines, *T. turgidum* (AABB) and *Ae. tauschii* (DD). We used two combinations of synthetic wheat parental lines for the comparison of overall gene expression profiles. Furthermore, we characterized genes that were specifically expressed during the course of allohexaploidization.

MATERIALS AND METHODS

Plant materials and RNA extraction We used two allopolyploid lines produced by the intergenic hybridization of different combinations of *T. turgidum* and *Ae. tauschii*. Allopolyploids of the first combination (C1) were produced by crossing *T. turgidum* L. ssp. *dicoccum* (Schränk Thell. (C1_AABB, accession no. KU124, AABB) with *Ae. tauschii* Coss. ssp. *strangulata* (C1_DD, accession no. KU2074, DD). The second combination (C2) was produced by crossing *T. turgidum* ssp. *durum* (C2_AABB, accession no. KU126, AABB) with *Ae. tauschii* ssp. *strangulata* (C2_DD, accession no. KU2075, DD). The new synthetic triploids (C1_S₀ and C2_S₀, ABD) were generated by embryo rescue (Inagaki et al., 1998), and the new synthetic hexaploids (C1_S₁ and C2_S₁, AABBDD) were produced by the self-pollination of colchicine-treated triploid plants (C1_S₀ and C2_S₀). The synthetic hexaploid lines KU-221-13 (produced in 1959) and KU-221-25 (produced in 1976) were provided by the National Bioresource Project (Laboratory of Crop Evolution, Graduate School of Agriculture, Kyoto University, Kyoto, Japan); the first generation of each of these synthetic hexaploid lines had been obtained by crossing KU124 (AABB) with KU2074 (DD) and KU126 (AABB) with KU2075 (DD), respectively (Tanaka, 1961). Subsequently, the new synthetic hexaploids of each combination were self-pollinated for 13 and four generations, respectively, at Kyoto University. The S₁₄ generation of KU-221-13 (C1_S₁₄) and the S₅ generation of KU-221-25 (C2_S₅) were

used as stable synthetic hexaploids in this study.

The plant materials used in the present study are shown in Supplementary Fig. S1. For natural hexaploids, *T. aestivum* cv. Chinese Spring (CS) and three Iranian landraces (NH1, PI622243; NH2, PI622233; and NH3, PI622268; Wang et al., 2013) were used. The Iranian landraces, obtained from the USDA ARS Germplasm Resources Information Network (GRIN; <http://www.ars-grin.gov/>), were sympatric with the paternal lines used in the present study, and both paternal lines were sympatric with *Ae. tauschii* lineage 2E, which is considered to be the main source of the D genome in southern Caspian Iran wheat (Wang et al., 2013). The sampled locations of these wheat lines are shown in Supplementary Fig. S2. The microarray data on the allohexaploids and their parental lines of C1 were from our previous study (Jung et al., 2014). 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.

Total RNA was isolated from biological triplicates of two-leaf seedlings (except for triploid plants, in which biological duplicates of two-leaf seedlings were used) with the Plant RNeasy Mini Kit (QIAGEN) according to the manufacturer's instructions.

RNA labeling and microarray hybridization First, 500 ng of RNA was labeled with cyanine 3-cytidine 5C-triphosphate (CTP) using the Agilent Low RNA Input Fluorescent Linear Amplification Kit (Agilent Technologies, Santa Clara, CA). Hybridization of samples to the 60-mer oligo-DNA microarray (Agilent Technologies), which contains 37,886 probes, and washings were performed as described by Kawaura et al. (2008). Microarray hybridizations were independently performed for two or more 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). The microarray data have been deposited into the NCBI Gene Expression Omnibus (accessions GSE64883 and GSE65262).

Microarray data analysis Raw fluorescence intensities (raw signal intensity > 1.0 threshold) were transformed to log₂ scale using GeneSpring GX software (v.12.6; Agilent Technologies). For normalization, transformed signals were shifted to the 75th percentile of all arrays and to the median of all probes. Quality control was performed as follows: (1) weakly hybridized probe spots were removed using the cut-off value of < 20% of the raw signal intensity of all microarray slides; (2) only probes flagged as 'detected' in at least one of the 10 samples were chosen; (3) repeatable probes were identified as those with a coefficient of variation of > 50%.

Mid-parent values (MPVs) between *T. turgidum*

(AABB) and *Ae. tauschii* (DD) for each combination were calculated by averaging the expression values across the hybridizations; the data set ratio of *T. turgidum* to *Ae. tauschii* was set at 1:1 *in silico* (Pumphrey et al., 2009; Rapp et al., 2009; Chagué et al., 2010). The moderated *t*-test ($P < 0.01$) was applied using GeneSpring GX (v.12.6) to detect expression differences between the two parental lines of each combination. To control for multiple testing errors, the false discovery rate (FDR) control of Benjamini and Hochberg (1995) was employed to control the FDR below the level of 0.05. One-way analysis of variance (ANOVA) was used for determining non-additive gene expression profiles with the Benjamini and Hochberg multiple testing correction procedure. For hierarchical clustering, GeneSpring GX (v.12.6), Cluster and Treeview software were used (<http://www.eisenlab.org/eisen/>; Eisen et al., 1998). All hierarchical clustering analyses were performed using Pearson's cosine distance metric and average linkage.

Gene ontology (GO) was analyzed using the original sequences of the microarray probes and GO terms were categorized as described previously (Jung et al., 2014). The GO terms were mapped to Plant GOSlim terms using AgBase GOSlimViewer (<http://www.agbase.msstate.edu/>). Enrichment analysis was conducted using the BiNGO tool in Cytoscape v.3.2.0 software (<http://www.cytoscape.org/>; Maere et al., 2005). For Kyoto encyclopedia of genes and genomes (KEGG) pathway analysis, the original sequences of microarray probes were assigned to the rice *TIGR* gene in a BLASTx search (E -value < $1E-20$). MSU gene 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 (<http://www.genome.jp/kegg/>) *Oryza sativa japonica* (RAPDB) databases.

RESULTS

Transcriptome differences between the parental lines in two independent crosses Global gene expression levels of the parental lines *T. turgidum* (C1_AABB, accession no. KU124, and C2_AABB, accession no. KU126) and *Ae. tauschii* (C1_DD, accession no. KU2074, and C2_DD, accession no. KU2075) were compared using an oligonucleotide microarray for common wheat consisting of 37,886 probes (Kawaura et al., 2008). In total, 18,248 genes that showed reliable expression among replicates were selected after stringent quality control using GeneSpring GX (v.12.6). A total of 7,308 (40.0%) differentially expressed genes (DEGs) were found between C1_AABB (KU124) and C1_DD (KU2074), and 7,229 (39.6%) DEGs were found between C2_AABB (KU126) and C2_DD (KU2075) (Table 1). This indicated that the tetraploid and diploid parents had distinct gene expression patterns. There were 4,633 DEGs (63.4% and 64%

Table 1. Number of common differentially expressed genes (DEGs) between progenitors of each combination

	Total	AABB > DD	AABB < DD
C1_AABB vs C1_DD	7308 (40.0%) ^a	3041 (41.6%) ^b	4267 (58.4%) ^b
C2_AABB vs C2_DD	7229 (39.6%) ^a	3111 (43.0%) ^b	4118 (57.0%) ^b
Common DEGs between (C1_AABB vs C1_DD) and (C2_AABB vs C2_DD)	4633 (63.4 and 64%) ^c	1875 (40.5%) ^b	2758 (59.5%) ^b

Differentially expressed genes were determined using a moderated *t*-test ($P < 0.01$) with false discovery rate $\alpha = 0.05$.

^a% of all detected genes (18,248).

^b% of DEGs between progenitors in each combination.

^c% of DEGs in combination 1 and 2.

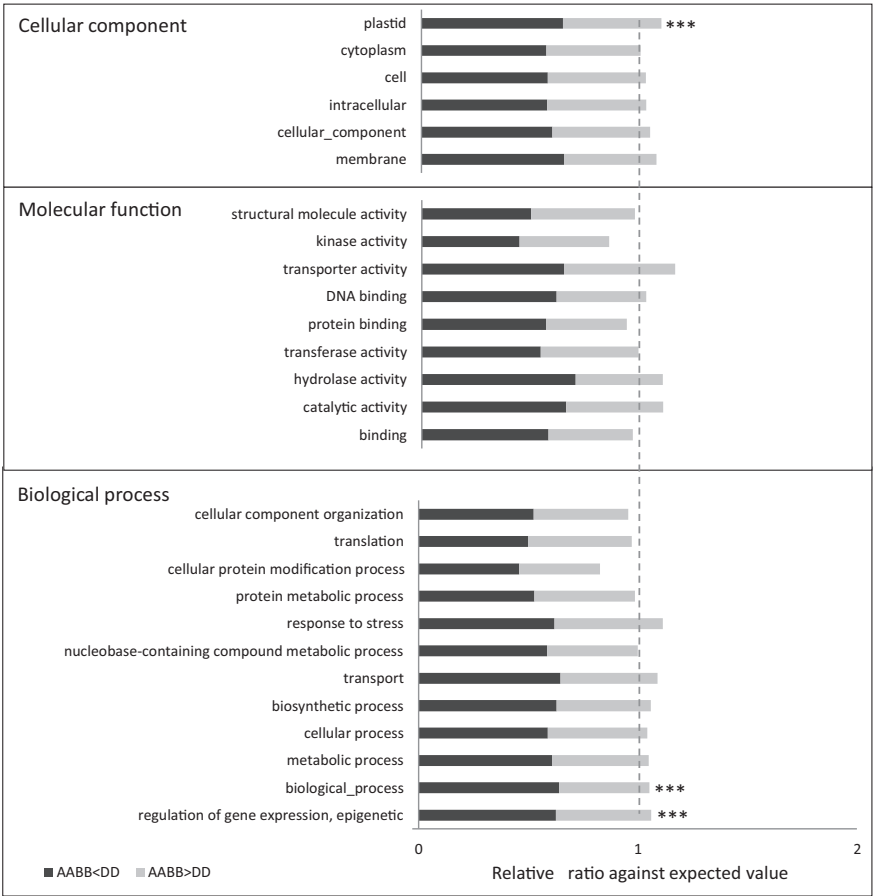


Fig. 1. Gene ontology (GOSlim) analysis of the differentially expressed genes common to the parental lines of each combination. GO terms are categorized into cellular component, molecular function and biological process. The *x*-axis represents the relative expression ratio of genes categorized by GO terms. The relative ratio was calculated as the ratio of genes categorized to a particular GO term divided by the ratio of all genes categorized to the same GO term in the microarray. Black bars indicate that the gene expression level is lower in *T. turgidum* (AABB) than in *Ae. tauschii* (DD), and gray bars indicate that the gene expression level is higher in *T. turgidum* (AABB) than in *Ae. tauschii* (DD). *** indicates significant enrichment of a particular GO term at $P < 0.001$ using false-discovery rates.

of DEGs in C1 and C2) that overlapped between the two combinations (Table 1), indicating that the majority of the changes in gene expression occurred in common in the independent crosses. Analysis of these common DEGs revealed that 1,875 (40.5%) were up-regulated in the maternal *T. turgidum* (AABB) lines, whereas 2,758 (59.5%) were up-regulated in the paternal *Ae. tauschii*

(DD) lines (Table 1). Next, we performed Gene Ontology (GOSlim) analysis of the 4,633 genes for each GO term category. Amongst all GO terms, the following were statistically enriched: ‘plastid’ in the cellular component category, and ‘biological process’ and ‘regulation of gene expression, epigenetic’ in the biological process category (Fig. 1). This result suggests that the expression of these

categories of common DEGs changed during the course of evolution of *T. turgidum* (AABB) and *Ae. tauschii* (DD).

Transcriptome differences during the course of allopolyploidization and genome stabilization To investigate the global relationship between gene expression patterns in the allopolyploids and parental lines of the two independent crosses, we conducted hierarchical clustering analysis (Eisen et al., 1998) based on the gene expression levels of these lines. The gene expression

patterns between the two independent crosses were closest to each other, indicating a similar gene expression pattern between crosses. A correlation dendrogram showed that the gene expression patterns of the new synthetic allohexaploids (C1_S₁ and C2_S₁) were more similar to those of the maternal *T. turgidum* (C1_AABB and C2_AABB) lines than were those of the subsequent allohexaploid generations (i.e., S₁₄ and S₅) (Fig. 2). No phenotypic differences were apparent between synthetic wheat of S₁ and S₁₄ or S₅ (Supplementary Fig. S1). Global

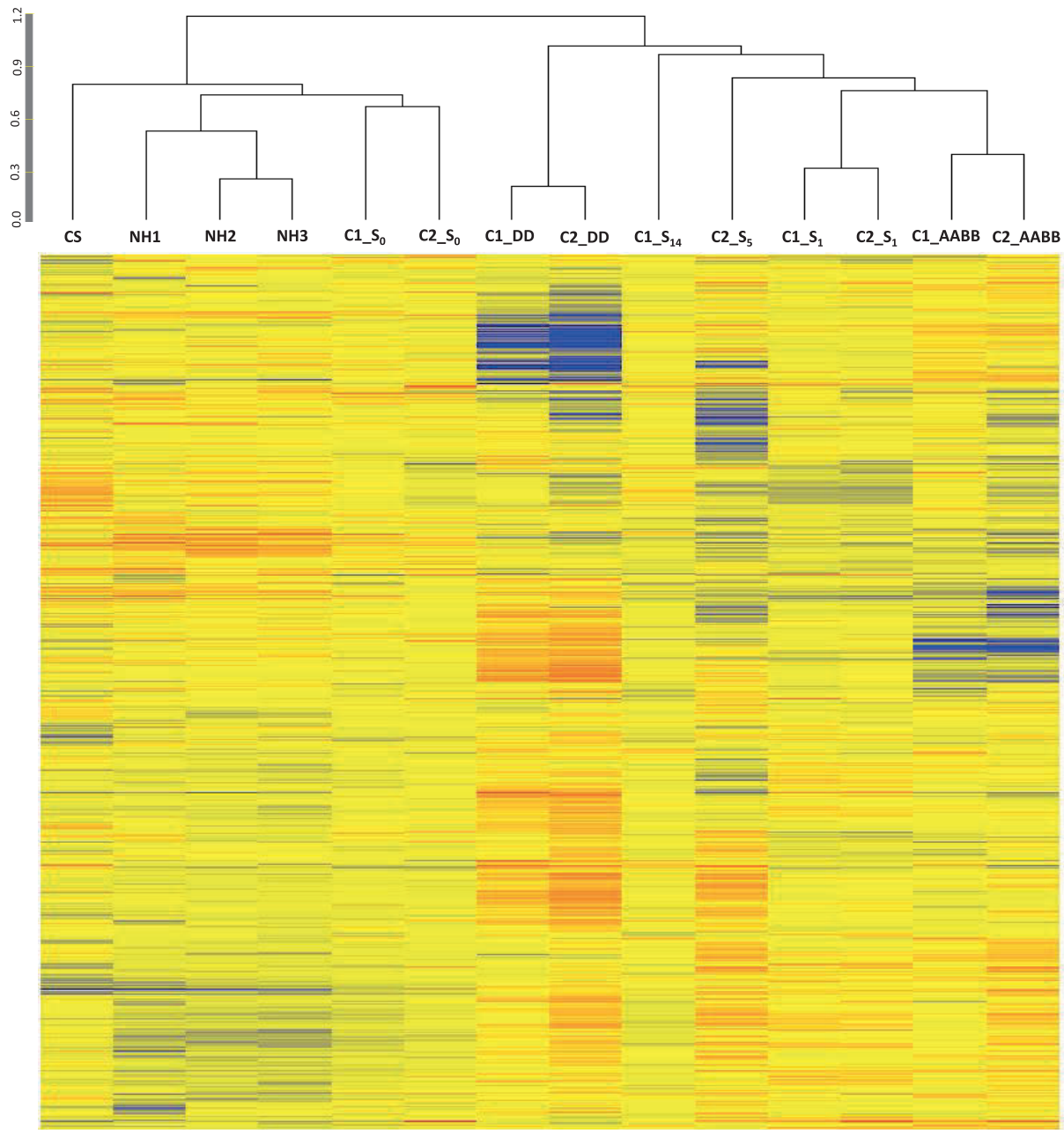


Fig. 2. Clustering analysis of gene expression patterns among synthetic allopolyploids and their parental lines in two independent combinations. The cluster dendrogram shows the global relationship of gene expression patterns among all of the synthetic allopolyploids and parental lines. Cluster analysis was performed using Pearson's cosine distance metric and average linkage.

gene expression patterns of embryo-rescued triploids were analyzed to investigate artificial effects due to the lack of endosperm through embryo culture. Interestingly, the gene expression patterns of the triploids (C1_S₀ and C2_S₀) were clustered into those of natural allohexaploid groups (CS, NH1-NH3), which were separate from those of the parental lines, including the paternal lines (DD). The clustering position of the triploids among all materials differed from that previously reported (Jung et al., 2014). This discrepancy may be due to how the polyploids were produced: the triploids in the previous study were produced from seeds (hereafter named Seed_C1_S₀ and Seed_C2_S₀), whereas the present triploids were rescued by embryo culture (EC_C1_S₀ and EC_C2_S₀). The relationships of gene expression patterns between triploids and their maternal parents are shown in Supplementary Fig. S3. The gene expression patterns of triploids grown from seeds correlated with those of their maternal tetraploids (Supplementary Fig. S3A), while triploids rescued by embryo culture exhibited no such correlation with either the maternal tetraploids (Supplementary Fig. S3B) or the triploids grown from seeds (Supplementary Fig. S3C).

Next, we compared the gene expression levels of the allopolyploids and MPVs. If allopolyploidization had no effect on gene expression levels, we would expect that any given gene would be expressed in the allopolyploid generation at the same level as the average of its parents (the MPV); these genes are referred to as additively expressed genes. In contrast, non-additively expressed genes (NEGs) display differences in expression between allopolyploids and MPVs. Non-additively expressed genes with a difference greater than 1.5-fold that of the MPVs were selected. The non-additive gene expression of three allopolyploid lines in the two independent crosses was profiled (Fig. 3A). In C1, a total of 3,898 NEGs were identified as follows: 2,741 (15.0%) were detected in the triploids (EC_C1_S₀); 2,645 (14.5%) were detected in the nascent allohexaploids (C1_S₁); and 1,421 (7.8%) were detected in the stable allohexaploids (C1_S₁₄). Among the 6,555 NEGs in C2, 4,948 (27.1%) were found in the triploids (EC_C2_S₀); 3,790 (20.8%) were found in the nascent allohexaploids (C2_S₁); and 3,539 (19.4%) were found in the stable allohexaploids (C2_S₅). These results indicated that the number of NEGs decreased along with allopolyploidization and genome stabilization.

Subsequently, the NEGs identified in the three allopolyploid lines were compared (Fig. 3, B and C). In C1, a total of 2,741 NEGs were detected in EC_C1_S₀, but 714 of these NEGs were not found in the other two generations, indicating that they were specific to the new hybridized triploid (Fig. 3B). Similarly, 833 NEGs in C1_S₁ were specific to generation S₁, whereas only 57 NEGs in C1_S₁₄ were specific to generation S₁₄. A total of 615 (15.8%) of 3,898 NEGs in C1 were commonly

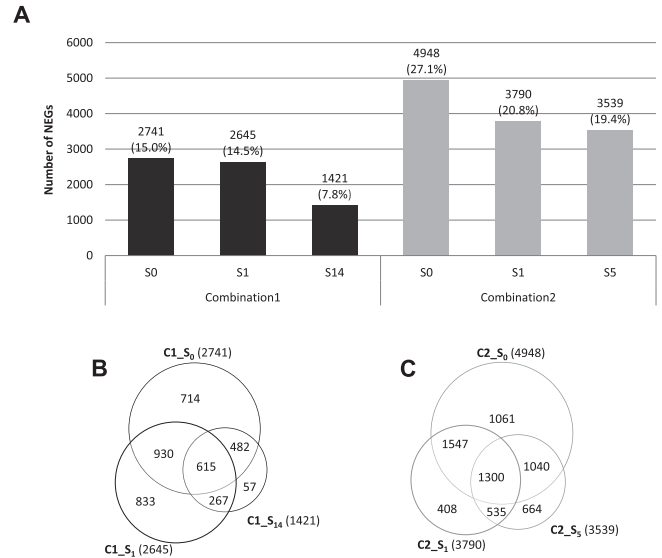


Fig. 3. Non-additively expressed genes of all allopolyploid lines in two combinations. (A) The number of NEGs in three allopolyploids in each of the two combinations. The numbers of NEGs in S₀, S₁ and S₁₄ of C1 (combination 1) are shown in black and those in S₀, S₁ and S₅ in C2 (combination 2) are shown in gray. The percentages of NEGs among all expressed genes are given in parentheses. (B, C) Venn diagrams of NEGs detected in the three generations of allopolyploids in combination 1 (B) and combination 2 (C). Overlapping regions indicate genes that were commonly detected in the different allopolyploids.

detected throughout the three allopolyploid lines. In C2, 1,061 NEGs in EC_C2_S₀ were specific to generation S₀, 408 NEGs in C2_S₁ were specific to generation S₁, and 664 NEGs in C2_S₅ were specific to generation S₅ (Fig. 3C). A total of 1,300 (19.8%) of 6,555 NEGs were found commonly throughout the three C2 allopolyploid lines. These results suggest that approximately 16% to 20% of NEGs were transmitted across generations throughout polyploidization and successive self-pollinations. On the other hand, GOSlim analysis of NEGs in the S₀, S₁ and S₅ generations is presented in Supplementary Table S1. Enriched categories after GOSlim analysis of NEGs changed from generation to generation, showing that NEGs were flexible among generations.

Heritable NEGs of two independent combinations

The gene expression patterns of the 615 and 1,300 heritable NEGs in the two combinations were classified into eight major groups based on comparisons with their MPVs (Fig. 4, A and B). These eight groups were designated as clusters A to H. The expression patterns and numbers of genes belonging to each cluster are shown in Fig. 4C. Genes in cluster A, comprising 90 and 65 genes in C1 and C2, respectively, showed down-regulation in all allopolyploid lines. In cluster B, 10 and 256 genes in C1 and C2, respectively, showed up-regulation specifically in the stable allohexaploid lines. In cluster C, eight and 28

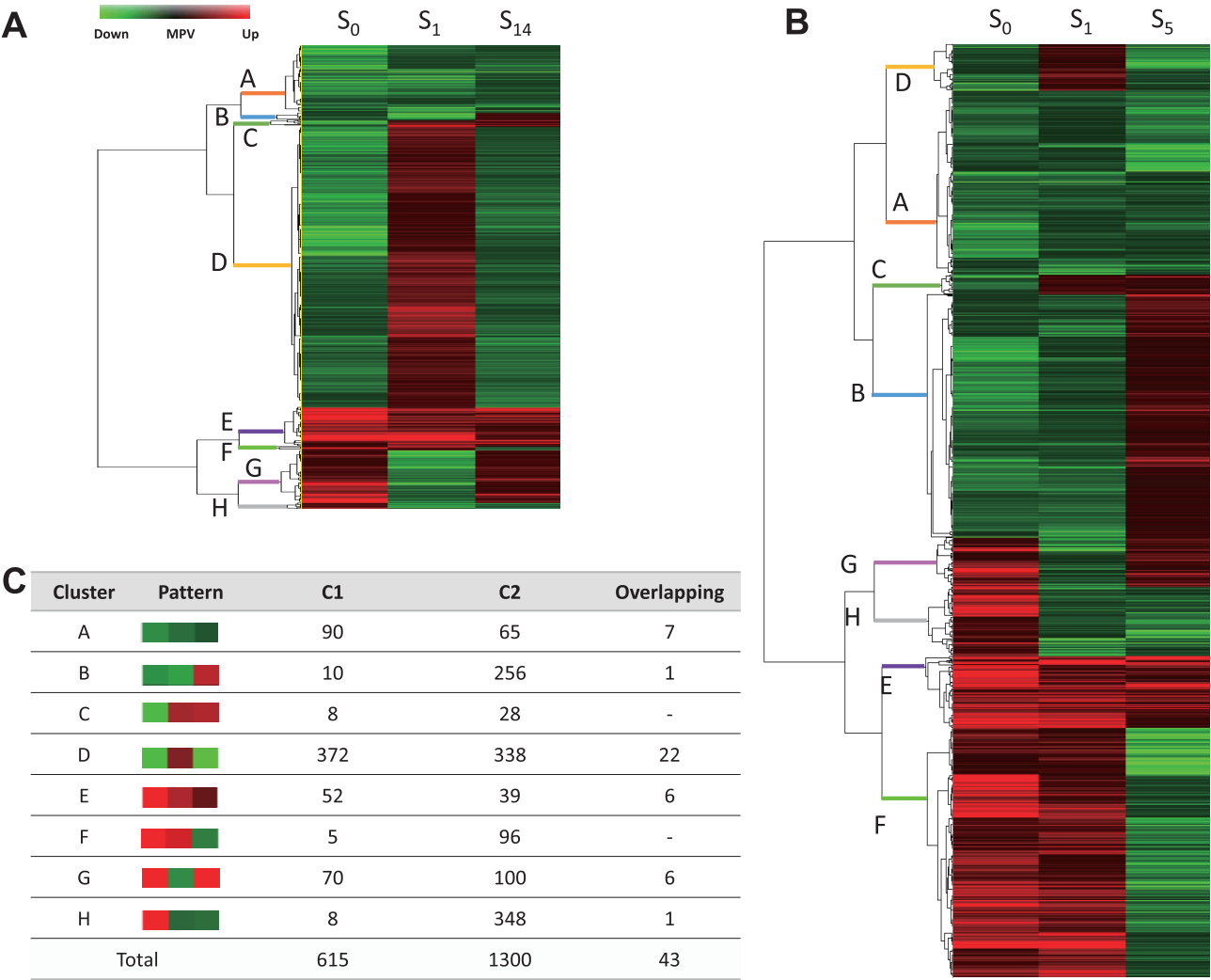


Fig. 4. Expression patterns of heritable NEGs in the three allopolyploid lines of two combinations. Hierarchical clustering of 615 and 1,300 heritable NEGs in the three allopolyploid lines of combination 1 (A) and combination 2 (B), respectively. In the color key at the top left corner, red represents a higher gene expression level than the MPV and green represents a lower gene expression level than the MPV. (C) Eight expression patterns were designated as groups A to H based on the hierarchical clustering of heritable NEGs in the two combinations.

genes in C1 and C2, respectively, showed down-regulation in the triploid lines. In cluster D, 372 and 338 genes in C1 and C2, respectively, showed up-regulation specifically in the new synthetic allohexaploid lines. In cluster E, 52 and 39 genes in C1 and C2, respectively, showed up-regulation in all allopolyploid lines. In cluster F, five and 96 genes in C1 and C2, respectively, showed down-regulation specifically in the stable allohexaploid lines. In cluster G, 70 and 100 genes in C1 and C2, respectively, showed down-regulation specifically in the new synthetic allohexaploid lines. In cluster H, eight and 348 genes in C1 and C2, respectively, showed up-regulation specifically in the triploid lines. These results indicate that the heritable NEGs were distinct between the two combinations of crosses. Among 615 and 1,300 heritable NEGs in the two combinations, 43 genes overlapped between

the two, indicating the presence of characteristic expression patterns during the course of allopolyploidization (Supplementary Table S2).

Shared NEGs regulate starch and sucrose metabolism To investigate the biological functions of the 43 NEGs that showed stable gene expression throughout generations in both lines, we used the KEGG database to analyze the genes. Nine genes were found to encode enzymes that possessed Enzyme Commission (EC) numbers, and they were all involved in metabolism: four were involved in starch and sucrose metabolism (starch phosphorylase (Os02g0330200, EC 2.4.1.1), soluble starch synthase (Os04g0624600, EC 2.4.1.21), cellulase (Os12g0428150, EC 3.2.1.4), and β -fructofuranosidase (Os02g0106100, EC 3.2.1.26)), three in the biosynthesis of

Table 2. Common heritable NEGs encoding enzymes associated with metabolism

KEGG pathway	Gene name	Functional annotation	EC number	Fold change against MPV for each combination					
				C1_S ₀	C1_S ₁	C1_S ₁₄	C2_S ₀	C2_S ₁	C2_S ₅
Starch and sucrose metabolism	wheat0130Contig15644	beta-fructofuranosidase	EC 3.2.1.26	6.5	8.9	1.6	15.0	7.1	1.6
	whok17f08	soluble starch synthase 1	EC 2.4.1.21	-1.7	4.6	-2.6	-2.6	1.7	-1.6
	wheat0130Contig3119	cellulase	EC 3.2.1.4	-2.2	-2.4	-1.6	-2.4	-2.1	-1.6
	MUGEST2003_23lib_Contig20208	starch phosphorylase	EC 2.4.1.1	-3.8	1.8	-1.6	-2.3	1.6	-5.2
Biosynthesis of secondary metabolites	ncbi_gi_32128941	beta-alanine ligase	EC 6.3.2.1	-2.9	-1.8	-1.6	-2.9	-1.6	-4.0
	wheat0130Contig11261	UDP-glucose 4-epimerase	EC 5.1.3.2	-3.4	3.6	-2.6	-2.5	3.4	-3.5
	wh22g14	asparagine synthase	EC 6.3.5.4	-18.6	-2.1	-2.3	-12.4	-1.6	-3.2
Purine metabolism	whh13n24	pyruvate kinase	EC 2.7.1.40	-2.9	2.2	-1.9	-3.0	2.3	-1.6
Glycerolipid metabolism	wheat0130Contig6657	triacylglycerol lipase	EC 3.1.1.3	5.8	2.6	5.1	6.2	3.8	6.4

secondary metabolites, and one each in purine and glycerolipid metabolism (Table 2).

DISCUSSION

Maternal expression dominance in the early stages of allohexaploidization We evaluated the effects of allopolyploidization and genome stabilization on gene regulation by comparing the allopolyploid lines of two combinations. The parental lines of two combinations of *T. turgidum* (AABB) and *Ae. tauschii* (DD) exhibited differential expression for 40.0% of the detected genes (Table 1 and Fig. 1). This frequency was equivalent to or slightly higher than those previously reported for wheat (Chagué et al., 2010; Qi et al., 2012) and other plant species such as *Gossypium* (Rapp et al., 2009) and *Arabidopsis* (Wang et al., 2006), suggesting that the differences in overall gene expression seen between the parental lines are due to intergenic divergence rather than ploidy level.

Next, we compared overall gene expression patterns among the allotriploids and nascent allohexaploids in two combinations with *T. turgidum* as the maternal parent and *Ae. tauschii* ssp. *strangulata* as the pollen parent (Fig. 2). Hierarchical clustering of expressed gene patterns in these lines showed that in the two combinations, those in the same generation were classified into closer clusters than those in different generations, indicating that allopolyploids displayed stage-specific expression patterns. Furthermore, it should be pointed out that nascent allohexaploids exhibited gene expression patterns that were more similar to those of their maternal allotetraploids than to those of pollen parents, and they constituted an outer group that was different from cultivated common wheats (Fig. 2); this occurred irrespective of whether they were produced by the crossing of unreduced gametes of tetraploids or by the colchicine treatment-induced chromosome doubling of hybridized

triploids (Tanaka, 1961). This maternal expression dominance has been reported in wheat (Qi et al., 2012; Li et al., 2014) and other plants such as *Spartina* species (Chelaifa et al., 2010).

In the present study, two nascent triploids (EC_C1_S₀ and EC_C2_S₀) clustered with the natural hexaploid wheat groups, rather than with the parental lines and nascent hexaploids (Fig. 2). This was contradictory to the results of our previous report (Jung et al., 2014), in which nascent triploids were closely related to the maternal tetraploid (AABB). Scatterplots showing the relationships of gene expression patterns between two S₀ lines, and between triploids and their maternal tetraploids (Supplementary Fig. S3), strongly supported the results of hierarchical clustering. This discrepancy can be explained by how the triploids were produced: the triploids in this study (EC_C1_S₀ and EC_C2_S₀) were produced by embryo rescue (in the absence of endosperm), while the triploids in the previous study (Seed_C1_S₀ and Seed_C2_S₀) were produced by hybridization and double fertilization. Furthermore, it should be pointed out that the gene expression profiles of the synthetic allohexaploids (C1_S₁ and C2_S₁), which were produced with the colchicine method from EC_C1_S₀ and EC_C2_S₀, resembled those of maternal lines rather than the natural hexaploids. It is notable that the D genome of *Ae. tauschii* was not included in the endosperm member of the triploids before fertilization. Additionally, maternal effects of gene expression patterns decreased with successive self-pollinations. These lines of evidence suggest that maternal dominance from double fertilization, which promotes maternal gene expression in the embryo of descendants, plays a key role in the stages of allopolyploidization and genome stabilization in plants (Nishiyama and Yabuno, 1978; Lafon-Placette and Köhler, 2014).

Non-additively expressed genes during the course of allopolyploidization and genome stabilization

Although the overall gene expression patterns of the nascent allohexaploids correlated relatively well with those of their maternal parent (AABB; Fig. 2) in the two combinations of allopolyploids, a number of NEGs were detected in each generation (Fig. 3). More NEGs were detected during allopolyploidization than after genome stabilization. This tendency was as previously reported (Jung et al., 2014), suggesting that genome shock (McClintock, 1984) occurred during intergenomic hybridization and the early stages of allopolyploidization, and that genome tuning among the three homoeologous genomes occurred through genome stabilization events over successive self-pollinations (Chagué et al., 2010; Chelaifa et al., 2010; Qi et al., 2012).

Persistent NEGs (Qi et al., 2012) are NEGs that are commonly detected during nascent allopolyploidization and they seem to characterize the phenotypes of allohexaploid wheat (Li et al., 2014); thus, heritable NEGs were selected from the NEGs detected in S_0 , S_1 and S_5/S_{14} of the two combinations of allopolyploid wheats (Fig. 4 and Supplementary Table S2). Out of the 43 selected persistent NEGs, four genes were related to starch and sucrose metabolism, three to the biosynthesis of secondary metabolites, and one each to purine and glycerolipid metabolism (Table 2). In *Arabidopsis thaliana*, ploidy level affected the growth vigor and starch content of the plant (Ni et al., 2009; Miller et al., 2012), and high-yield and fast-growing plants showed a higher concentration of sugars (Sun et al., 2012). Phenomena found in allopolyploid wheat seem to be similar to those in *Arabidopsis*. These data suggest that sugar and/or starch metabolism is a key pathway for the establishment of allopolyploidization and genome stabilization, and for promoting hybrid vigor in hexaploid wheat. Although the exact molecular mechanisms remain unknown, allopolyploidy-induced sudden reunification of different genomes may involve gene expression modulation through transcription factors (Ng et al., 2011), transposable elements (Yaakov and Kashkush, 2011; Zhao et al., 2011), epigenetic remodeling and/or small interfering RNAs (Ni et al., 2009; Ha et al., 2011; Zhao et al., 2011; Ng et al., 2012; Shen et al., 2012). Therefore, further studies are necessary to identify the precise mechanisms of regulation and the functions of NEGs, as well as the importance of allopolyploidization in the hybrid vigor of allohexaploid wheat.

This work was partially supported by Grants-in-Aid for scientific research in the priority area of "Comparative Genomics" from the Ministry of Education, Culture, Sports, Science and Technology of Japan. This is contribution no. 1018 from the Kihara Institute for Biological Research, Yokohama City University.

REFERENCES

- Akhunova, A. R., Matniyazov, R. T., Liang, H., and Akhunov, E. D. (2010) Homoeolog-specific transcriptional bias in allopolyploid wheat. *BMC Genomics* **11**, 505.
- 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.
- Brenchley, R., Spannagl, M., Pfeifer, M., Barker, G. L. A., D'Amore, R., Allen, A. M., McKenzie, N., Kramer, M., Kerhornou, A., Bolser, D., et al. (2012) Analysis of the bread wheat genome using whole-genome shotgun sequencing. *Nature* **491**, 705–710.
- 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.
- Chelaifa, H., Monnier, A., and Ainouche, M. (2010) Transcriptional changes following recent natural hybridization and allopolyploidy in the salt marsh species *Spartina x townsendii* and *Spartina anglica* (Poaceae). *New Phytol.* **186**, 161–174.
- Chelaifa, H., Chagué, V., Chalabi, S., Mestiri, I., Arnaud, D., Deffains, D., Lu, Y., Belcram, H., Huteau, V., Chiquet, J., et al. (2013) Prevalence of gene expression additivity in genetically stable wheat allohexaploids. *New Phytol.* **197**, 730–736.
- Cheng, F., Wu, J., Fang, L., Sun, S. L., Liu, B., Lin, K., Bonnema, G., and Wang, X. W. (2012) Biased gene fractionation and dominant gene expression among the subgenomes of *Brassica rapa*. *PLoS ONE* **7**, e36442.
- Dubcovsky, J., and Dvorak, J. (2007) Genome plasticity a key factor in the success of polyploid wheat under domestication. *Science* **316**, 1862–1866.
- Dvorak, J., Deal, K. R., Luo, M.-C., You, F. M., von Borstel, K., and Dehghani, H. (2012) The origin of spelt and free-threshing hexaploid wheat. *J. Hered.* **103**, 426–441.
- Eisen, M. B., Spellman, P. T., Brower, P. O., and Botstein, D. (1998) Cluster analysis and display of genome wide expression patterns. *Proc. Natl. Acad. Sci. USA* **95**, 14863–14868.
- Feldman, M., Levy, A. A., Fahima, T., and Korol, A. (2012) Genomic asymmetry in allopolyploid plants: wheat as a model. *J. Exp. Bot.* **63**, 5045–5059.
- Feuillet, C., Langridge, P., and Waugh, R. (2008) Cereal breeding takes a walk on the wild side. *Trends Genet.* **24**, 24–32.
- 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.
- Hatano, H., Mizuno, N., Matsuda, R., Shitsukawa, N., Park, P., and Takumi, S. (2012) Dysfunction of mitotic cell division at shoot apices triggered severe growth abortion in interspecific hybrids between tetraploid wheat and *Aegilops tauschii*. *New Phytol.* **194**, 1143–1154.
- 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.
- Jung, Y., Kawaura, K., Mishina, K., Sakuma, S., Kishii, M., and Ogihara, Y. (2014) Changes in genome-wide gene expression during allopolyploidization and genome stabilization in hexaploid wheat. *Genes Genet. Syst.* **89**, 215–225.
- 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.

- Kihara, H. (1944) Discovery of the DD-analyser, one of the ancestors of *Triticum vulgare*. *Agric. Hortic.* **19**, 13–14.
- Kihara, H., Okamoto, M., Ikegami, M., Tabushi, J., Suemoto, H., and Yamane, Y. (1950) Morphology and fertility of five new synthesized hexaploid wheats. *Seiken Zihō* **4**, 127–140 (Japanese with English summary).
- Lafon-Placette, C., and Köhler, C. (2014) Embryo and endosperm, partners in seed development. *Curr. Opin. Plant Biol.* **17**, 64–69.
- Leitch, A. R., and Leitch, I. J. (2008) Genomic plasticity and the diversity of polyploidy plants. *Science* **320**, 481–483.
- Li, A., Liu, D., Wu, J., Zhao, X., Hao, M., Geng, S., Yan, J., Jiang, X., Zhang, L., Wu, J., et al. (2014) mRNA and small RNA transcriptomes reveal insights into dynamic homeolog regulation of allopolyploid heterosis in nascent hexaploid wheat. *Plant Cell* **26**, 1878–1900.
- Maere, S., Heymans, K., and Kuiper, M. (2005) *BiNGO*: a Cytoscape plugin to assess overrepresentation of Gene Ontology categories in biological networks. *Bioinformatics* **21**, 3448–3449.
- Mallet, J. (2007) Hybrid speciation. *Nature* **446**, 279–283.
- McClintock, B. (1984) The significance of responses of the genome to challenge. *Science* **226**, 792–801.
- McFadden, E. S., and Sears, E. R. (1944) The artificial synthesis of *Triticum spelta*. *Rec. Genet. Soc. Am.* **13**, 26–27.
- 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.
- 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.
- Nishijima, R., Iehisa, J. C. M., Matsuoka, Y., and Takumi, S. (2014) The cuticular wax inhibitor locus *Iw2* in wild diploid wheat *Aegilops tauschii*: phenotypic survey, genetic analysis, and implications for the evolution of common wheat. *BMC Plant Biol.* **14**, 246.
- Nishiyama, I., and Yabuno, T. (1978) Causal relationships between the polar nuclei in double fertilization and interspecific cross-incompatibility in *Avena*. *Cytologia* **43**, 453–466.
- 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.
- Pont, C., Murat, F., Guizard, S., Flores, R., Foucrier, S., Bidet, Y., Quraishi, U. M., Alaux, M., Doležel, J., Fahima, T., et al. (2013) Wheat syntenome unveils new evidences of contrasted evolutionary plasticity between paleo- and neoduplicated subgenomes. *Plant J.* **76**, 1030–1044.
- Pumphrey, M., Bai, J., Laudencia-Chingcuanco, D., Anderson, O., and Gill, B. S. (2009) Nonadditive expression of homeologous 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.
- Rieseberg, L. H., and Willis, J. H. (2007) Plant speciation. *Science* **317**, 910–914.
- Schnable, J. C., Springer, N. M., and Freeling, M. (2011) Differentiation of the maize subgenomes by genome dominance and both ancient and ongoing gene loss. *Proc. Natl. Acad. Sci. USA* **108**, 4069–4074.
- Shaked, H., Kashkush, K., Ozkan, H., Feldman, M., and Levy, A. A. (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.
- 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.
- Soltis, D. E., Visger, C. J., and Soltis, P. S. (2014) The polyploidy revolution then...and now: Stebbins revisited. *Am. J. Bot.* **101**, 1057–1078.
- Soltis, P. S., and Soltis, D. E. (2009) The role of hybridization in plant speciation. *Annu. Rev. Plant Biol.* **60**, 561–588.
- Sun, F., Suen, P. K., Zhang, Y., Liang, C., Carrie, C., Whelan, J., Ward, J. L., Hawkins, N. D., Jiang, L., and Lim, B. L. (2012) A dual-targeted purple acid phosphatase in *Arabidopsis thaliana* moderates carbon metabolism and its overexpression leads to faster plant growth and higher seed yield. *New Phytol.* **194**, 206–219.
- Tanaka, M. (1961) New amphidiploids, synthesized 6x-wheats, derived from Emmer wheat x *Aegilops squarrosa*. *Wheat Inf. Ser.* **12**, 11.
- Wang, J., Luo, M. C., Chen, Z., You, F. M., Wei, Y., Zheng, Y., and Dvorak, J. (2013) *Aegilops tauschii* single nucleotide polymorphisms shed light on the origins of wheat D-genome genetic diversity and pinpoint the geographic origin of hexaploid wheat. *New Phytol.* **198**, 925–937.
- Wang, J., Tian, L., Lee, H. S., Wei, N. E., Jiang, H., Watson, B., Madlung, A., Osborn, T. C., Doerge, R. W., Comai, L., et al. (2006) Genomewide nonadditive gene regulation in *Arabidopsis* allotetraploids. *Genetics* **172**, 507–517.
- 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.
- Yakimowski, S. B., and Rieseberg, L. H. (2014) The role of homoploid hybridization in evolution: a century of studies synthesizing genetics and ecology. *Am. J. Bot.* **101**, 1247–1258.
- Zhang, H. K., Bian, Y., Gou, X. W., Dong, Y. Z., Rustgi, S., Zhang, B. J., Xu, C. M., Li, N., Qi, B., Han, F. P., et al. (2013) Intrinsic karyotype stability and gene copy number variations may have laid the foundation for tetraploid wheat formation. *Proc. Natl. Acad. Sci. USA* **110**, 19466–19471.
- Zhang, L., Liu, D., Zheng, Y., Yan, Z., Dai, S., Li, Y., Jiang, Q., Ye, Y., and Yen, Y. (2010) Frequent occurrence of unreduced gametes in *Triticum turgidum*-*Aegilops tauschii* hybrids. *Euphytica* **172**, 285–294.
- 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.