Theor Appl Genet (2010) 120:341–353
DOI 10.1007/s00122-009-1113-3

ORIGINAL PAPER

Gene expression analysis at the intersection of ploidy and hybridity in maize

Nicole C. Riddle · Hongmei Jiang · Lingling An · R. W. Doerge · James A. Birchler

Received: 5 April 2009 / Accepted: 13 July 2009 / Published online: 6 August 2009

Abstract Heterosis and polyploidy are two important aspects of plant evolution. To examine these issues, we conducted a global gene expression study of a maize ploidy series as well as a set of tetraploid inbred and hybrid lines. This gene expression analysis complements an earlier phenotypic study of these same materials. We find that ploidy change affects a large fraction of the genome, albeit at low levels; gene expression changes rarely exceed 2-fold and are typically not statistically significant. The most common gene expression profile we detected is greater than linear increase from monoploid to diploid, and reductions from diploid to triploid and from triploid to tetraploid, a trend that mirrors plant stature. When examining heterosis in tetraploid maize lines, we found a large fraction of the genome impacted but the majority of changes were not statistically significant at 2-fold or less. Non-additive expression was common in the hybrids, and the extent of non-additivity increased both in number and magnitude from duplex to quadruplex hybrids. Overall, we find that gene expression trends mirror observations from the phenotypic studies; however, obvious mechanistic connections remain unknown.

Communicated by M. Kearsey.

Contribution to the special issue "Heterosis in Plants".

Electronic supplementary material The online version of this article (doi:10.1007/s00122-009-1113-3) contains supplementary material, which is available to authorized users.

N. C. Riddle · J. A. Birchler (✉)
Division of Biological Sciences, University of Missouri,
117 Tucker Hall, Columbia, MO 65211, USA
e-mail: birchlerj@missouri.edu

H. Jiang · L. An · R. W. Doerge
Department of Statistics, Purdue University,
West Lafayette, IN 47907, USA

Present Address:
N. C. Riddle
Department of Biology, Washington University,
St. Louis, MO 63130, USA

Present Address:
H. Jiang
Department of Statistics, Northwestern University,
Evanston, IL 60208, USA

Present Address:
L. An
Department of Agricultural and Biosystems Engineering,
University of Arizona, Tucson, AZ 85721, USA

Introduction

Polyploidy has played an important role in plant evolution having occurred repeatedly followed by diploidization, which happens by gene loss with retention of selected gene categories (Freeling and Thomas 2006). An interesting issue concerns the reasons for the success of polyploids, which may intersect with the phenomenon of hybrid vigor—the superior performance of hybrid individuals over more inbred parents.

Previous studies have characterized the phenotypes of a maize ploidy series as well as tetraploid inbreds and hybrids (Riddle et al. 2006; Riddle and Birchler 2008). In a 1–4x ploidy series of inbred lines, the monoploid shows a reduced but not highly detrimental stature compared to diploids. Increases in ploidy above the diploid using inbred derivatives always produced a decline in stature and vigor. However, with hybrids at the tetraploid level, duplex hybrids exhibit heterosis more or less comparable to

diploids but in quadruplex tetraploid hybrids with four potential alleles at each locus, even greater heterosis is found for several characteristics (Riddle and Birchler 2008; Levings et al. 1967; Sockness and Dudley 1989a, b). This phenomenon of progressive heterosis parallels results from other species (Busbice and Wilsie 1966; Groose et al. 1989; Mok and Peloquin 1975; Bingham et al. 1994).

Recent studies have compared gene expression in diploid inbreds and hybrids (Swanson-Wagner et al. 2006; Guo et al. 2006; Stupar and Springer 2006; Huang et al. 2006; Uzarowska et al. 2007; Stupar et al. 2008; Hoecker et al. 2008). In this study, we examine global patterns of gene expression in a ploidy series and tetraploid hybrids as a complement to the phenotypic studies. These patterns were found to differ between lines, ploidy levels and hybrids but without obvious correlations to phenotype. Nevertheless, the most common trend of gene expression in the ploidy series involved an increase from monoploid to diploid followed by a progressive decline to the triploid and tetraploid, which mirrors the plant stature trend. In tetraploid hybrids, non-additive expression of some genes was observed in hybrids with a greater number and of increasing magnitude in the quadruplex hybrids, again mirroring the plant stature trend but without obvious connections between the phenotype and gene expression.

Materials and methods

Plant material

Plant material for the B73 ploidy series is derived from the standard diploid inbred line B73 of *Zea mays*. The four genotypes included in this study are monoploid B73 (1x), diploid B73 (2x, the progenitor to the other three genotypes), triploid B73 (3x), and tetraploids B73 (4x). The monoploid B73 material was generated using the method described (Auger et al. 2004). Triploid B73 plants were generated by treatment of tassels with trifluralin (Kato 1997, 1999a, b) to generate diploid sperm that produce triploid zygotes upon fertilization. Tetraploids were derived from diploid B73 (Kato and Birchler 2006).

The A188, W22, and Oh43 tetraploid lines were derived from the standard diploid lines (Kato and Birchler 2006). The tetraploids inbred lines were crossed to generate single-cross F1 hybrids (A188/Oh43 [AO] and B73/W22 [BW]). The F1 hybrids were crossed to each other to create the double-cross quadruplex hybrid A188/Oh443/W22/B73 (AOWB; A188/Oh43 × W22/B73).

All plants for tissue collection were grown in the Sears greenhouse at the University of Missouri, Columbia with 16-h light per day. ProMix BX General Purpose Growing Medium (Premier Horticulture) supplemented with iron sulfate, DynaGreen^{TM} 12-12-12 (Hummert) fertilizer and greensand was used and additional liquid fertilizer (a 0.3% solution of Peter's general purpose 20-20-20 fertilizer; Hummert).

RNA isolation

Leaf tissue was collected by harvesting adult, fully expanded leaves from 10 plants per genotype between 2 and 3 p.m. (the plants were approximately 1 month old). The pooled tissue was frozen in liquid nitrogen and stored at −80°C until processing. Frozen leaf tissue, excluding midribs, was ground to a fine powder in liquid nitrogen using a mortar and pestle. Total RNA was isolated from the leaf tissue using Trizol (Invitrogen) following the manufacturer's recommendation with minor modifications. In particular, a 1-h incubation on ice was performed after the addition of isopropanol to the sample, prior to the precipitation of the RNA. The RNA was resuspended in water and stored at −80°C until processing. For each genotype, two independent RNA samples (A and B) were prepared.

Microarray preparation

Oligonucleotide microarrays were obtained from the Maize Oligonucleotide Array Project at the University of Arizona. An array set consisting of two slides contains 57,452 70mer oligonucleotides. Before hybridization, the microarray slides were post-processed according to the manufacturer's recommendations. In brief, they were rehydrated by holding them DNA side down over a 50°C water bath for 10 s, followed by snap-drying on a 65°C heat block for 10 s. These two steps were repeated a total of four times. Subsequently, the DNA was immobilized on the glass slide in a UV crosslinker at 65 mJ. Excess oligonucleotides were removed by washing the slides for 5 min in 1% sodium dodecyl sulfate (SDS), while traces of SDS were removed by dipping the slides in water. Before drying the arrays with a short spin in a slide centrifuge, the arrays were washed with 95% ethanol for 3 min.

Probe preparation

An indirect labeling method using amino-allyl-dUTP followed by coupling to Cy3/Cy5 was carried out according to a modified protocol (Catts et al. 2005). To generate cDNA, 20 µg of total RNA were incubated in 1× First Strand Buffer (supplied with the RT (reverse transcriptase) enzyme from Invitrogen), 5-mM DTT (dithiothreitol), 3.75 µM of an anchored oligo dT primer ($dT_{20}VN$; IDT) for 5 min at 65°C, followed by another 5-min incubation at 42°C (concentrations given refer to the concentration in the complete reaction after the addition of the reverse transcriptase

enzyme). dGTP, dATP, and dCTP were added to a final concentration of 0.5 mM, dTTP to a final concentration of 0.16 mM, and amino-allyl-dUTP to a concentration of 0.338 mM (dNTPs: Invitrogen; aa-dUPT: Sigma). Finally, 400 U of Superscript III RT (Invitrogen) were added. The reactions were incubated overnight at 42°C. The RNA was destroyed by adding 4 µl 50-mM ethylenediaminetetraacetic acid (EDTA), 2 µl 10-M NaOH (sodium hydroxide) and incubated the samples at 65°C for 20 min. The pH was neutralized by the addition of 4 µl 5-M acetic acid. The RT reactions were purified using the QIAquick PCR purification kit (QIAGEN) according to the manufacturer's recommendations with the following modifications: (1) two 70% ethanol washes were used to wash the column instead of the recommended wash with buffer PE. (2) The cDNA was eluted from the column using 30 µl of water twice. Using a vacuum concentrator, the sample was dried to approximately 2–3 µl. Cy3 and Cy5 fluorescent dyes are coupled to the amino-allyl-labeled cDNA by incubating the cDNA with 9 µl 0.1-M $NaHCO_3$, pH 9.0 and 2-µl Cy3 or Cy5 *N*-hydroxy succinimide ester [dye-packs from Amersham; each vial resuspended in 18-µl DMSO (dimethyl sulfoxide)]. This reaction was incubated at room temperature in the dark for 45 min. The coupling reactions were washed using the QIAquick PCR purification kit (QIAGEN) according to the manufacturer's recommendations with the same modifications given above. Each Cy3/Cy5 labeled sample was then dried to 2–5 µl in a vacuum concentrator.

Hybridization

The hybridization mix consisted of DIG Easy Hyb (Roche) supplemented with 500-µg/ml yeast tRNA and 500-µg/ml salmon sperm DNA for blocking. Two Cy3 and two Cy5 labeled probes were combined with 120 µl of this mix to be used on one slide set. The hybridization mix was heated to 65°C for 5 min and then cooled to room temperature before loading onto the microarray slide. The array surface was covered with a coverslip (Lifterslip, Erie Scientific), the array was moved into a hybridization chamber (Corning) and incubated at 37°C for approximately 15 h. After hybridization, the coverslip was removed by immersing the array into 1× sodium chloride/sodium citrate buffer (SSC) at room temperature. The slide was washed thrice 10–15 min in 1× SSC, 0.1% SDS at 50°C (10% SDS, 20× SSC from Ambion). After the last wash, residual SDS was removed by a quick rinse in 1× SSC at room temperature. The slide was dried by centrifugation in a slide centrifuge.

Scanning

An Axon GenePix 4000B scanner was used to scan all microarrays. Unless otherwise noted, three scans were performed per slide, all at 100% laser power: (1) A scan where the Cy3/Cy5 ratio had been adjusted to approximately 1 using a prescan at low resolution by altering the photo multiplier tube (PMT) gain. (2) A scan where the gain settings were constant among all slides (500 for 532/Cy3 and 600 for 635/Cy5). (3) A scan where the PMT was determined by the automatic setting of the scanner, with a pixel saturation level of 0.005%. For experiments 1 and 2, two additional scans were performed: a scan with the Cy3/Cy5 ratio adjusted to 1 at 33% laser power and a high-intensity scan which had the PMT gain of scan two from above increased by 100 for either channel. GenePix software was used to grid the arrays.

Experimental design

A loop experimental design consisting of dye swaps was employed to study gene expression differences of the various genotypes (illustrated in Fig. 1) allowing comparisons such as monoploid B73, diploid B73, triploid B73, and tetraploid B73. The statistical details of these analyses are in the Supplemental Materials. Sixteen slide sets were utilized for each dye swap experiment. Specifically, four slide sets of sample 1A were labeled with Cy3. Similarly, sample 2A was labeled with Cy5. The samples and dyes were then switched and another four slide sets were employed. Samples 1B and 2B were handled in a similar manner, leading to a total of 16 slide sets per comparison. Samples with A and B here refer to the duplicate RNA preps per genotype. Due to problems with individual arrays, for the comparison between monoploid and diploid B73, only 14 slide sets were used for the analysis. For all other comparisons, 16 slide sets were used.

The loop design (Fig. 1) allowed comparisons between tetraploid inbreds A188 and Oh43 or B73 and W22 to their

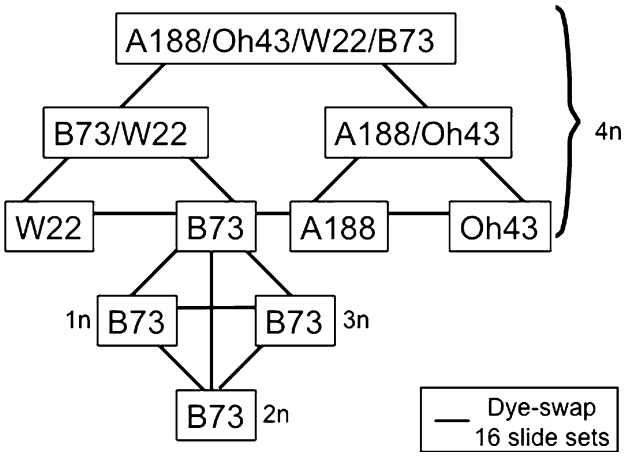

Fig. 1 Experimental design. The comparisons of gene expression levels between B73 lines of varying ploidy and inbreds with hybrid tetraploid lines was carried out in the loop design shown in the diagram. Each line connecting two genotypes represents dye-swap experiments with 16 slide sets

respective F1 hybrid. Sixteen slide sets were utilized for each of the three comparisons in each loop. In the A188/Oh43 loop, there were two instances of data loss. For the comparison between Oh43 and the A188/Oh43 hybrid, one set of data for slide B with scan setting 1 was missing due to a corrupted data file; its dye swap partner was excluded to achieve a balanced design. In the experiments comparing A188 and the A188/Oh43 hybrid, two slides A datasets were excluded due to technical problems. In the second loop, for the comparison between the B73 and W22 with scan setting 3 only seven dye swaps were analyzed (for slide B). Similarly, in the comparison between B73 and the B73/W22 hybrid only seven dye swaps were analyzed for all scan settings (for slide A).

For the remaining comparisons the following analyses did not include the full eight dye swap datasets: in the comparison between B73/W22 hybrid to the double-cross hybrid, seven dye swaps were analyzed for slide B; in the comparison between B73 diploids and B73 tetraploids, seven dye swaps were analyzed for slide B; and in the comparison between B73 monoploids and triploids, seven datasets were analyzed for both slides A and B.

Statistical analysis

Statistical analyses were performed as described in Supplemental Materials. Briefly, background-subtracted median signal intensity values for each of the 54,492 features were log2 transformed prior to analysis. Two types of analysis of variance (ANOVA) were carried out, one assuming a common variance for all features and one using a feature-specific variance. Statistical significance using a multiple comparison correction was determined in two ways, a Bonferroni correction with $\alpha = 0.05$, and false discovery rate (FDR) controlling procedure (Benjamini and Hochberg 1995), again with $\alpha = 0.05$. Both approaches control for false positives (i.e., Type I error, gene is statistically differentially expressed when in fact it is not) when performing multiple tests. The Bonferroni method is used to control the family wise error rate, the probability of having any false positives, while the FDR controlling procedure is used to control the FDR, the expected proportion of false positives among all the rejected hypotheses. The Bonferroni method is more conservative in the sense that the statistically significant results detected by the Bonferroni method are also statistically significant under the FDR criterion.

Results

Ploidy series

The steady state mRNA levels in monoploid, diploid, triploid, and tetraploids of the B73 maize line was examined using oligonucleotide microarrays (for summary statistics, see Supplemental Table 1). While the differences in ploidy affect many plant characteristics and also developmental timing, we chose fully expanded adult leaves as the source of RNA, because they can be readily harvested in large quantities. Six comparisons were carried out: monoploid to diploid, monoploid to triploid, monoploid to tetraploid, diploid to triploid, diploid to tetraploid, and triploid to tetraploids. Four of these comparisons were performed directly with arrays, while the remaining two were estimated based on the loop design (see Supplemental Materials).

Comparison of monoploid and diploid

The first comparison performed was between monoploid and diploid plants (Fig. 2a). Using an FDR multiple testing correction procedure, we detected 9,132 genes with significant expression differences between leaf tissue from monoploid and diploid plants. Using a more stringent Bonferroni correction yielded 775 significant differentially expressed genes, all of which were detected by the FDR procedure. Of these statistically significant genes 5,190 exhibited a reduced expression in the monoploid tissue compared to the diploid tissue, while 3,942 genes increased (FDR). Only 26 of the 5,190 statistically significant genes exhibiting a decrease in expression were changed by more than 2-fold. As a point of reference, among the statistically significant genes the overall average level of reduced expression was 1.27-fold, and 2.37-fold for a majority of the 26 significant genes that displayed a reduction in expression but exhibited no apparent functional relationship. Interestingly, a larger number of statistically significant genes increased by more than 2-fold when compared to 2-fold down modulated genes despite there being a much larger number of genes being reduced overall. The average level of expression increase was 1.25-fold, which is very similar to the overall level of reduction. However, 40 statistically significant genes had an increased expression of more than 2-fold (on average 2.44-fold). Although statistically significant expression differences were documented in comparisons between monoploids and diploids, the fold differences are quite small, almost all of them less than 2-fold, and a variety of gene classes were involved.

Comparison of diploid and triploid

When investigating expression differences between diploid and triploid plants, a large number of statistically significant changes were detected (Fig. 2b). Based on a FDR procedure, 15,413 genes exhibited significant differential expression; this number was reduced to 3,623 genes when the more conservative Bonferroni procedure was employed. Only a small fraction of these genes was down

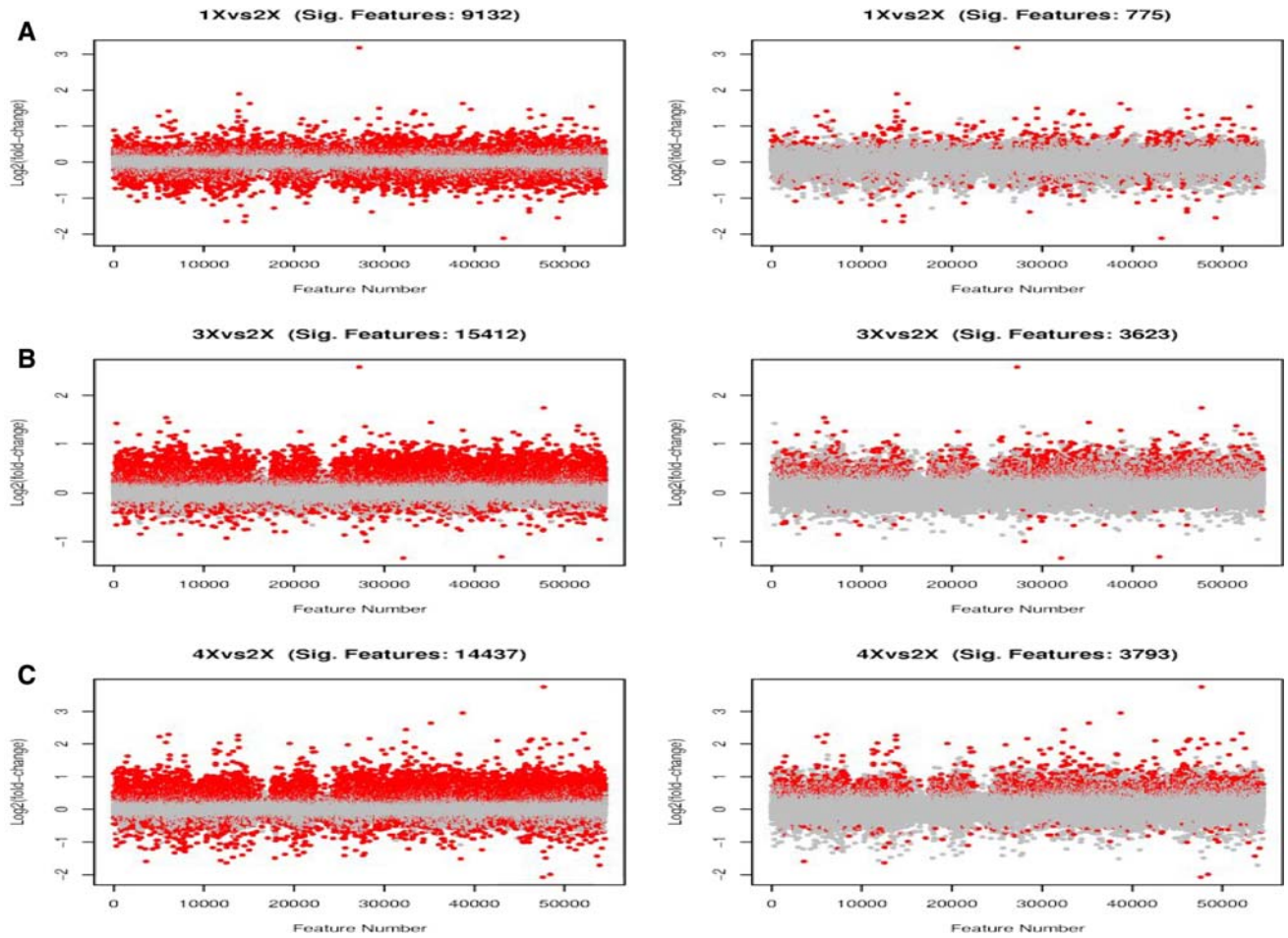

Fig. 2 Gene expression differences observed in the B73 ploidy series. For each comparison (1x vs. 2x, 3x vs. 2x, 4x vs. 2x) genes (features on the microarray, *X*-axis) are plotted against the log2 of the fold change observed between the two genotypes (log2(fold change); *Y*-axis). Genes with significant up- or down-regulation are shown in *red*, non-significant genes are in *gray* (data from scan 3). **a** Monoploid–diploid comparison, **b** triploid–diploid comparison, **c** tetraploid–diploid comparison; *left*: FDR criterion, *right*: Bonferroni criterion

modulated in triploid B73 plants compared to diploid plants, namely 667 or 4% (FDR). Among the statistically significant genes the average level of reduction was 1.26-fold. Of these, only three were reduced more than 2-fold: a papain-like cysteine peptidase, a C4 sterol methyl oxidase, and a subtilisin/chymotrypsin inhibitor. On the other hand, 14,745 genes were significantly increased (FDR), and 51 of these were increased more than 2-fold. The average increase was 1.30-fold overall. The 51 most increased genes had an average 2.29-fold change. Thus, when comparing triploid gene expression to that of diploid plants, the most common occurrence was that of greater expression. This change was less than 2-fold for 99% of genes that were identified.

Comparison of diploid and tetraploid

Differential expression was tested between diploid and tetraploid plants (Fig. 2c). A total of 3,793 genes were detected with significantly altered gene expression between the diploid and tetraploid maize lines using the more conservative Bonferroni multiple comparison correction, while 14,437 genes were detected with the FDR correction. Similar to the diploid–triploid comparison, 1,339 genes were significantly reduced (FDR), 80 of which were modulated more than 2-fold. The average level of reduction was 1.41-fold overall, and 2.33-fold among the most highly reduced genes. Thus, the magnitude of change increased from the triploid to the tetraploid. The average level of change among the 13,098 significantly increased genes (FDR) was 1.46-fold, a level that increased to 2.36-fold among the 479 genes with more than a 2-fold change. As noticed for the reduced genes, we saw an increase in the magnitude of changed expression when comparing results from the triploid–diploid and tetraploid–diploid. This trend was evident not only in the increase of average fold change, but also in the increase in the number of genes with fold changes greater than 2.

Ploidy series—composite analysis

Given the expression level changes in the B73 ploidy series, we examined data for any patterns in expression response to ploidy change. The results are illustrated in the Venn diagram shown in Supplemental Figure S4. There were 2,729 genes that were identified as significantly changed at all three irregular ploidy levels by the FDR, and 143 such genes according to the more conservative Bonferroni procedure (Supplemental Figure S4). It was also evident from this simple analysis that triploid and tetraploid plants share the majority of changed genes. On the other hand, the monoploid plants exhibited a quite distinct set, and only a small fraction of the genes changed in the monoploids was also detected as significant in the plants of higher ploidy.

The relationship between the genes altered in all three unnatural ploidy levels is further illustrated in Fig. 3. Plotting the fold-change values for monoploid against the fold change for either triploid or tetraploid plants resulted in a very similar slope. The correlation between the two measures is relatively low at 0.43 for the monoploid–tetraploid comparison and 0.47 for the monoploid–triploid comparison. In contrast, the correlation between fold change measures for triploid and tetraploid B73 was much higher at $R^2 = 0.70$, and the slope was significantly greater for this comparison than for the previous two (compare Fig. 3a–c). This same relationship was also evident when considering genes significant by the FDR criteria; a steeper slope and higher correlation were seen for the triploid and tetraploid gene expression measures than for the other two comparisons (Supplemental Figure S5). A positive slope indicates that if a gene was increased in the triploids compared to the diploids, it was increased more so in the tetraploids.

Ploidy series—hierarchical clustering

Another question we addressed was whether there were certain sets of genes that exhibited a similar response to changes in ploidy. To achieve this goal, we performed a hierarchical clustering analysis on the genes that showed more than minimal expression in all four B73 lines examined (51,225 genes; minimal expression was defined as a signal intensity of 1(=$\log_2(2)$) after background subtraction). Clustering was carried out in two steps. First, genes were grouped according to their overall expression response to ploidy change, with three possible responses to each stepwise change in ploidy, up (U), down (D), or no change (E). Thus, $3 \times 3 \times 3 = 27$ groups were identified; for example "UDU" grouped the genes that increase in expression from monoploids to diploids, reduced from diploids to triploids, but increased from triploids to tetraploids. Within each of the 27 groups, we performed hierarchical clustering to identify subgroups. The distribution of genes among the 27 main groups and the number of subclusters are summarized in Supplementary Table 2. Overall, the majority of genes increased in expression going from monoploid to diploid (28,397), and the remaining genes were roughly evenly divided between the reduced (10,768) and unchanged class (12,060). This distribution contrasted with that of the other two steps in the ploidy series, diploid to triploid and triploid to tetraploid. In both latter cases, the largest class of genes clustered into the decreased class (23,326 and 23,839 genes, respectively). The remaining genes again divided fairly evenly into the increased class (15,941 and 14,974) and the unchanged class (11,958 and 12,412 genes).

Figure 4 illustrates that some classes were clearly under- and over-represented. The largest group of genes adhered to the following pattern: increase from monoploid to diploid, and decrease from diploid to triploid and from triploid to tetraploid. The 8,424 genes fell into seven clusters, with three of the clusters containing the majority of genes. In addition to this predominant group, two groups in particular had very few members, DDD and UUU. These two groups represented genes that show a simple linear relationship with ploidy: either gene expression decreased with increasing ploidy (DDD) or gene expression increased with increasing ploidy. The clustering data indicate that there appear to be few genes that exhibited a linear relationship between expression levels and ploidy.

Hybrid versus inbred tetraploids

To gain insight into the impact of hybridity on gene expression in tetraploids, we investigated two different sets of tetraploid hybrids and their parental lines. The analysis was conducted using a loop design, shown in Fig. 1. The results of these analyses for the A188/Oh43 experiment are summarized in Supplementary Table 3 and in Fig. 5. A comparison of tetraploid inbred lines A188 and Oh43 revealed in excess of 75% of the genes assayed could be detected as showing a significant difference between the two lines. Most of these genes showed small differences in expression levels. A total of 49,801 of the genes exhibited a lower expression level in A188 than in Oh43 tetraploids, while 4,691 genes were more highly expressed in A188. This bias in gene expression remained apparent when focusing on genes with significantly altered gene expression, although not necessarily as pronounced. The differences between the two lines were extensive but only a small fraction of genes showed differences in excess of 2-fold.

Comparing the A188/Oh43 hybrid to A188 revealed differences for 75% of the transcriptome. A total of 50,671 genes showed lower expression in A188 than in the hybrid,

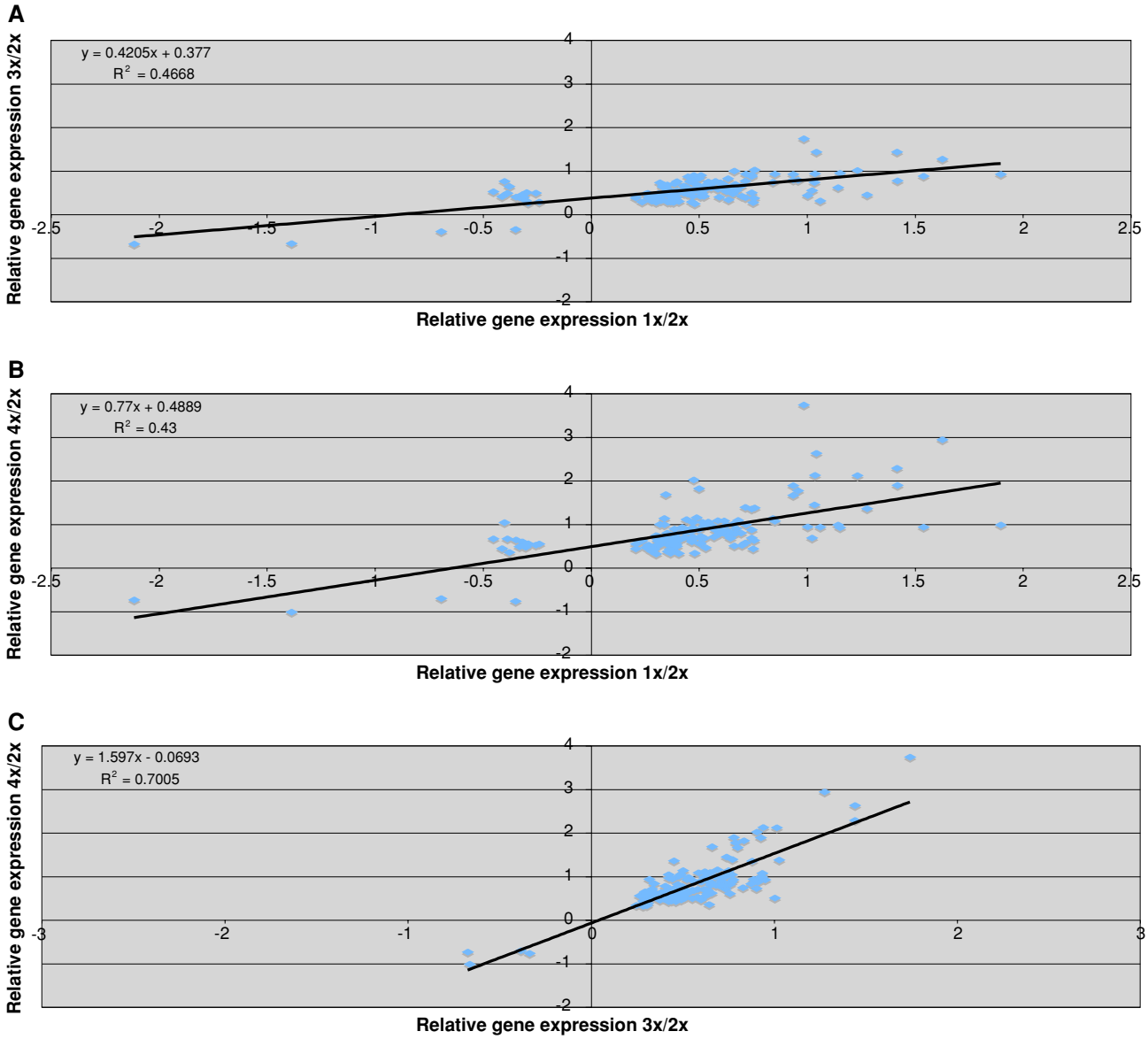

Fig. 3 Genes significantly different from the diploid norm in tetraploids and triploids are highly correlated, while the genes with altered expression in the monoploid show a distinct pattern. Included in this graph are only genes with significantly altered expression levels in all three abnormal ploidy levels compared to the diploid (scan 3, Bonferroni correction). In each section, the relative gene expression of one ploidy compared to diploid (log2 (fold change); *X*-axis) is plotted against that of a second ploidy compared to diploid (*Y*-axis) to gain an understanding of how the differences observed relate to one another. At the *top left* of each graph, the equation for the linear trendline as well as the R^2 is given. **a** *X*-axis: log2 (fold change) monoploid/diploid; *Y*-axis: log2 (fold change) triploid/diploid. **b** *X*-axis: log2 (fold change) monoploid/diploid; *Y*-axis: log2 (fold change) tetraploid/diploid. **c** *X*-axis: log2 (fold change) triploid/diploid; *Y*-axis: log2 (fold change) tetraploid/diploid

while only 3,821 genes showed higher expression in A188. These numbers were very similar to what was observed in the A188–Oh43 comparison.

In contrast to the previous two comparisons, only a smaller number of genes were differentially expressed between Oh43 and the A188/Oh43 hybrid. The gene expression pattern of the hybrid resembled the Oh43 parent much more than the A188 parent. The average level of increase among these highly significant genes was greater than 2-fold, while the average reduction was slightly higher at greater than 2.4-fold.

The second tetraploid hybrid comparison involved the B73 and W22 inbreds. The results of these analyses are summarized in Supplementary Table 4. B73 and W22 are genetically quite distinct, with large numbers of polymorphisms characterizing each line. Overall, we found less differences in gene expression between B73 and W22, than were identified in the A188–Oh43 comparison (compare

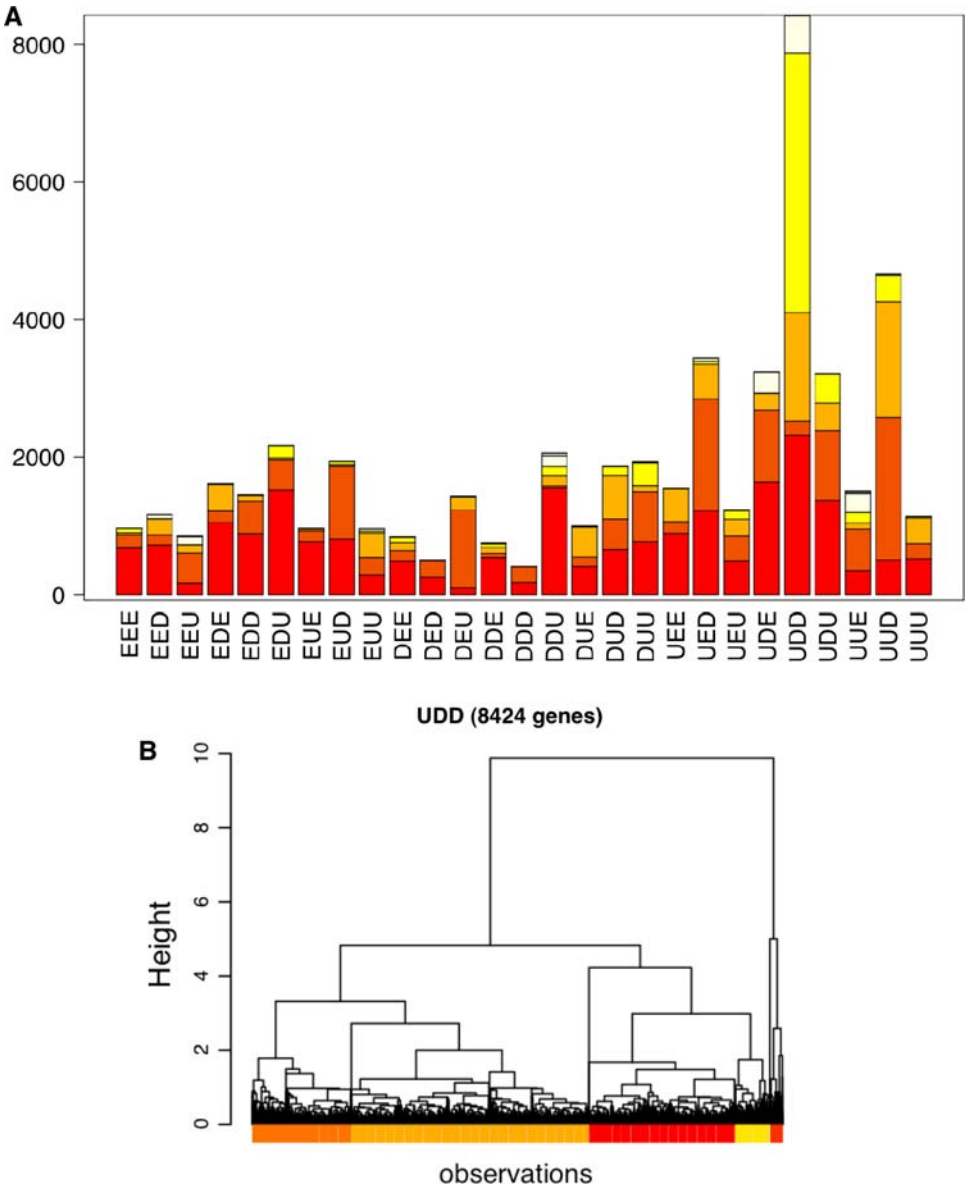

Fig. 4 Gene clusters identified in the B73 ploidy series. **a** Overview of all clusters and subclusters. Each section (*X*-axis) represents one major cluster, while the number of features within the major clusters is illustrated by its height (*Y*-axis). The number of subclusters and their size are illustrated by *color* within any given bar. **b** Details of the hierarchical clustering results for the UDD group. Results of the hierarchical clustering carried out within the major cluster UDD (expression levels go up from 1x to 2x, and down from 2x to 3x and 3x to 4x). The genes are arranged according to their subcluster membership along the *X*-axis (see *color diagram*)

Supplementary Table 3 to Supplementary Table 4). Overall, the differences were small (see Fig. 6). A total of 27,624 genes were increased in B73 compared to W22, while 27,654 genes showed the opposite pattern. Using both the FDR and the Bonferroni multiple comparison corrections, 1,422 were expressed at a higher level in B73, while 1,233 were expressed at higher levels in W22. Thus, the highly significant genes exhibited the same pattern noticed overall; there was no strong bias in the genes detected as differentially expressed toward one inbred line or the other.

There were 116 genes reduced more than 2-fold in W22 compared to B73, with an average modulation of approximately 2.4-fold. Among the genes increased in W22, there were 99 genes with a modulation of more than 2-fold. Their average increase was approximately 2.4-fold, identical to the average reduction observed for this experiment. The genes differentially expressed in B73 and W22 tetraploid inbreds were functionally diverse and most gene expression differences were quite small.

The number of genes detected as significantly different between the B73 and hybrid genotypes was very similar to the number detected when comparing B73 and W22 (Supplemental Table 4). There was a relatively equal distribution of up- and down-modulation. Among these statistically significant differentially expressed (FDR and Bonferroni) genes, 615 were more highly expressed in the B73 inbred line, while the remaining 493 genes were more highly expressed in the hybrid. Only 8 genes showed an increase of expression in the hybrid in excess of 2-fold (on average $\sim$ 2.3-fold).

Lastly, we examined the relationship between gene expression in the second parent, W22, and the B73/W22

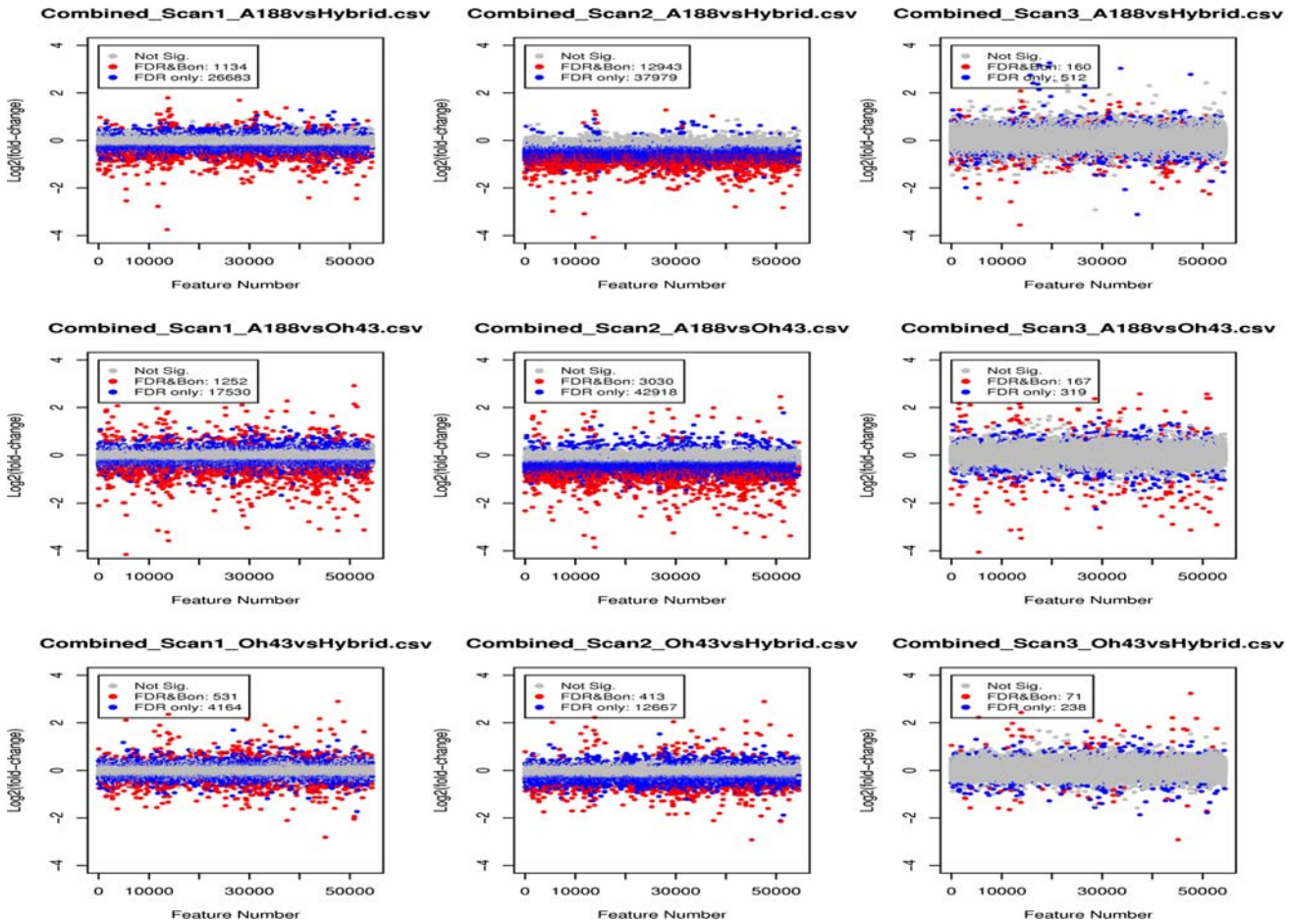

Fig. 5 Gene expression changes observed in the A188–Oh43 comparison. For each comparison (A188 vs. A188/Oh43 hybrid, A188 vs. Oh43, and Oh43 vs. A188/Oh43 hybrid), genes (features on the microarray, *X*-axis) are plotted against the log2 of the fold change observed between the two genotypes (log2 (fold change); *Y*-axis). Genes with significant expression modulation are shown in either *red* (meet both FDR and Bonferroni criterion) or *blue* (meet FDR criteria only); non-significant genes are in *gray*. The three columns show data from scans 1, 2, and 3. *Top*: A188 vs. A188/Oh43 hybrid, *middle*: A188 vs. Oh43, *bottom*: Oh43 vs. A188/Oh43 hybrid

hybrid. The number of differentially expressed genes detected was very similar to what was found in the comparison of the two inbred parent lines and in the other parent/hybrid comparison—up to ~20% of the transcripts assayed were detected as significant depending on the analysis method (Supplemental Table 4; Fig. 6). This result indicated that in terms of gene expression the B73/W22 hybrid did not favor one of the parents—as was observed in our experiment with A188 and Oh43—but rather that it was equally dissimilar to both parents.

Overall, gene expression levels were equally likely to be either up- or down-modulated in W22 compared to the B73/W22 hybrid. Focusing on just those genes that were detected as significantly different between W22 and the B73/W22 hybrids with the loop analysis and the Bonferroni correction, we identified 1,607 genes. Of these, 866 genes were increased in W22 compared to the hybrid, while 742 genes were significantly decreased. Among the 68 genes, whose expression was increased by more than 2-fold in W22, the average level of modulation was ~2.4-fold. Of the 743 reduced genes, 28 genes were modulated by more than 2-fold, with an average level of increase of ~2.4-fold.

Dominance patterns of gene expression in A188/Oh43 and B73/W22 hybrids

Next, we investigated the relationship between the gene expression levels in the inbred tetraploid lines and their F1 hybrid offspring. The simplest model predicts that gene expression in the hybrid should be equivalent to the average between the two parents. However, this model does not hold true for morphological characters due to heterosis. Thus, we compared gene expression levels of the A188/Oh43 hybrid to the average or midparent value between the A188 and Oh43 lines, and the B73/W22 hybrid to the average or midparent value between the B73 and W22 lines. We found that 2,165 genes were significantly different from the midparent value in the A188/Oh43 hybrid. In the B73/

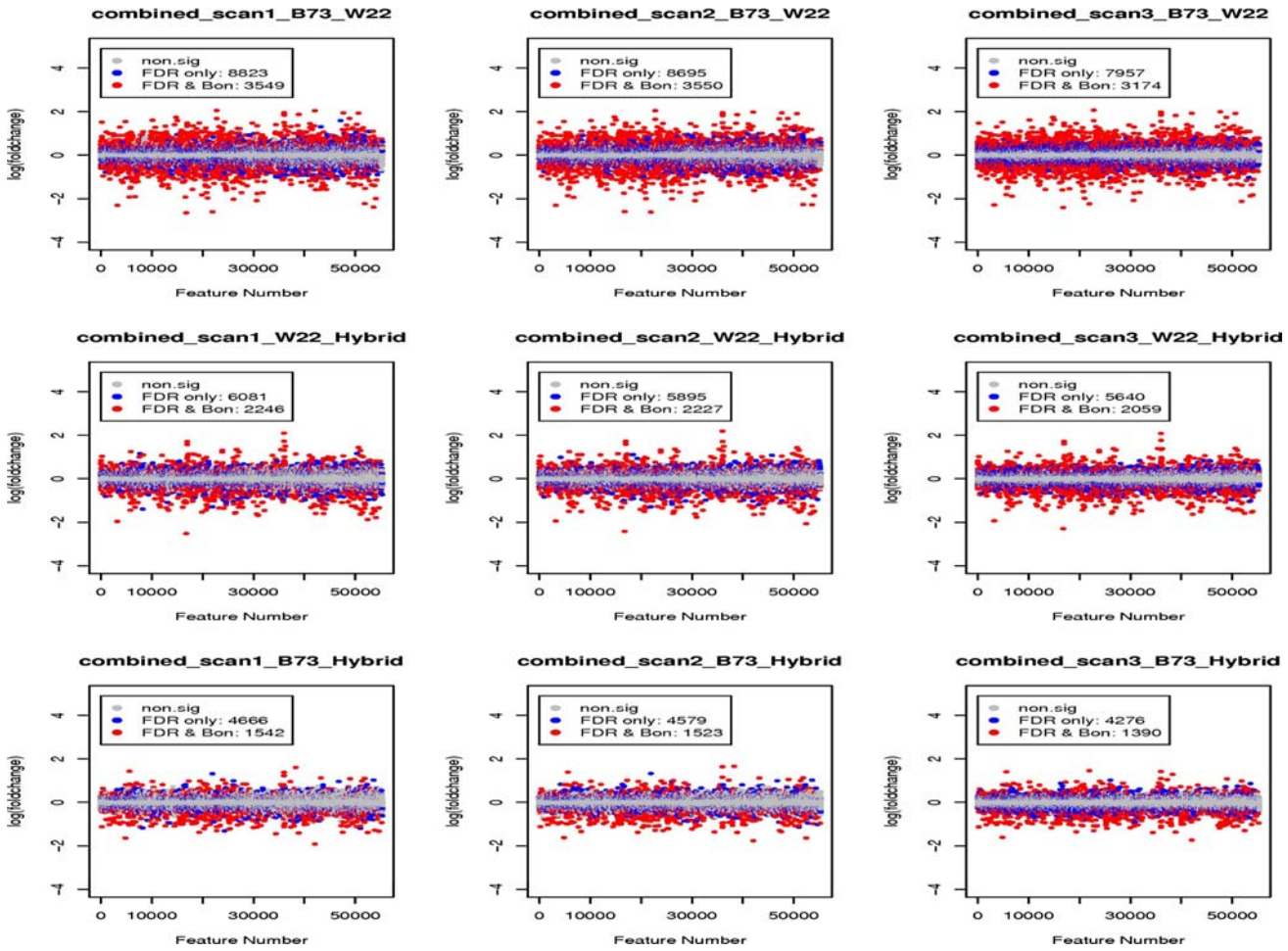

Fig. 6 Gene expression changes observed in the B73–W22 comparison. For each comparison (B73 vs. B73/W22 hybrid, B73 vs. W22, and W22 vs. B73/W22 hybrid), genes (features on the microarray, *X*-axis) are plotted against the log2 of the fold change observed between the two genotypes (log2 (fold change); *Y*-axis). Genes with significant modulation are shown in either *red* (meet both FDR and Bonferroni criterion) or *blue* (meet FDR criteria only), nonsignificant genes are in *gray*. The three columns show data from scans 1, 2, and 3. *Top*: W22 vs. B73, *middle*: B73/W22 hybrid vs. W22, *bottom*: B73/W22 hybrid vs. B73

W22 hybrid, we found 7,426 genes that differed from the midparent. A total of 780 genes were detected in both comparisons.

An analysis of the genes exhibiting non-additive gene expression levels in the hybrids is shown in Supplemental Table 5. The genes were classified into four categories; first into "above" or "below", indicating if the gene expression level in the hybrid was higher or lower than the midparent value, and then into "dominant" or "transgressive" depending on if the expression level for the hybrid was more extreme than that of both parents (transgressive) or not (dominant). Interestingly, for both hybrids investigated, transgressive behavior prevailed among the genes that differed significantly from the midparent value. In the A188/Oh43 hybrid, 770 genes exhibited transgressive gene expression in excess of the parental values, while 1,289 genes were expressed significantly below the level of both the A188 and Oh43 parents. In contrast, only 37 and 69 genes, respectively, demonstrated strictly dominant behavior above or below the midparent value. Similar results were obtained from the analysis of the B73/W22 hybrid. Most genes significantly different from the midparent values exhibited transgressive gene expression patterns: 3,060 above the level of both parents, 3,489 below. A total of 452 genes were strictly dominant above the midparent value, 245 were dominant below the midparent value in the B73/W22 hybrid.

Given the genes detected as significantly different from the midparent in the A188/Oh43 and the B73/W22 hybrids, we determined if any of these genes were affected in both hybrids, and if so, if the effects seen were similar. As noted above, 780 transcripts were identified in both hybrid analyses. As shown in Fig. 7, there was no clear relationship detected between the gene expression differences in relation to the respective midparent exhibited in one hybrid versus the other. When only the genes detected as significant were

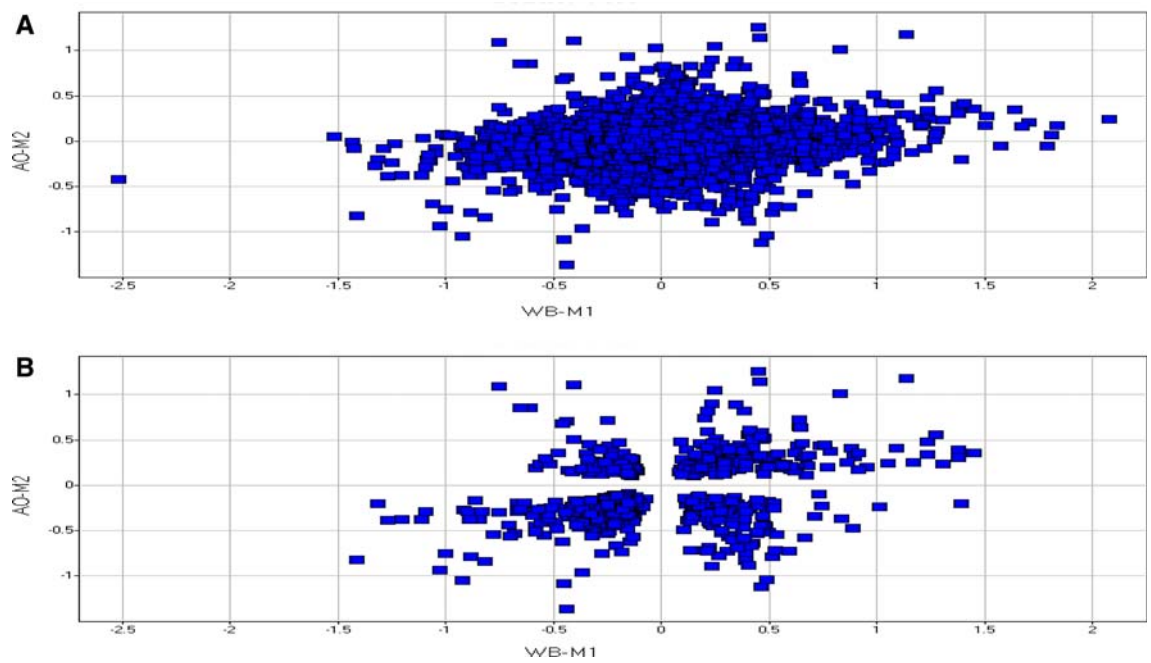

Fig. 7 Relationship between gene expression differences of the A188/Oh43 and B73/W22 hybrids and their respective midparents. **a** No correlation is detected when all genes are considered. *X*-axis: B73/W22 (log2 (fold change) of the hybrid compared to the midparent); *Y*-axis: A188/Oh43 (log2 (fold change) of the hybrid compared to the midparent). **b** Using the same axes definitions as in (**a**), two groups of genes are distinguishable when only genes with significant deviation from the midparent are considered, those with a positive correlation in the two hybrids, and those with a negative correlation

plotted (Fig. 7b), it became obvious that there were some genes that showed the same gene expression trends in both hybrids (and thus a positive correlation) and a second group of gene with opposite trends (leading to a negative correlation). However, the genes did not fall into specific groups that collectively showed a positive or negative correlation when comparing the two hybrids.

Dominance patterns of gene expression in the quadruplex hybrid A188/Oh43/W22/B73

In addition to the duplex hybrids A188/Oh43 and B73/W22, our study also included the quadruplex hybrid A188/Oh43/W22/B73 derived from a cross between two duplex hybrids. Comparing this quadruplex hybrid to the duplex hybrids as well as to the inbred tetraploid parent lines allowed for the assessment of the question of gene expression changes with increased allelic diversity.

First, we compared the gene expression levels observed in the A188/Oh43/W22/B73 (AOWB) quadruplex hybrid to four different midparent values: M_1, the midparent between B73 and W22; M_2, the midparent between A188 and Oh43; M_3, the midparent between the two duplex hybrids A188/Oh43 (AO) and B73/W22 (BW); and M_0, the midparent among all four tetraploid inbred lines, A188, Oh43, B73, and W22. A summary of the results is shown in Supplemental Table 6. Overall, fairly similar numbers of genes were detected as significantly different from the midparent. As in the analyses of the duplex hybrids, there was a clear trend that transgressive gene expression was equally likely to be increased or decreased compared to the respective midparents.

Combining data from the four quadruplex analyses with the data from the two duplex analyses, we identified a total of 73 genes with consistent transgressive behavior in all cases. Of these genes, 7 showed gene expression levels higher than of any parent, while 66 showed levels below all parental values. Of the 7 genes showing consistent transgressive gene expression above the parental level, only 2 were annotated, a *Zeon1* (a maize retrotransposon) gag protein, and a beta-carotene hydroxylase. Among the 66 transcripts with expression levels below parental values, 18 transcripts lacked annotation information, and a large cluster of 16 transcripts were annotated as ribosomal proteins. We also identified elongation factors (3), plastidic cysteine synthases (3), pathogenesis-related proteins (2), chloroplast proteins (2), an allyl alcohol dehydrogenase, a homocysteine *S*-methyltransferase, lipoxygenase, a peroxidase, a protease, and several other proteins.

Lastly, we compared gene expression patterns in duplex and quadruplex hybrids. Compared to their respective midparent, both types of hybrids exhibited cases of transgressive gene expression, i.e., gene expression that was more extreme, either positive or negative, than either direct

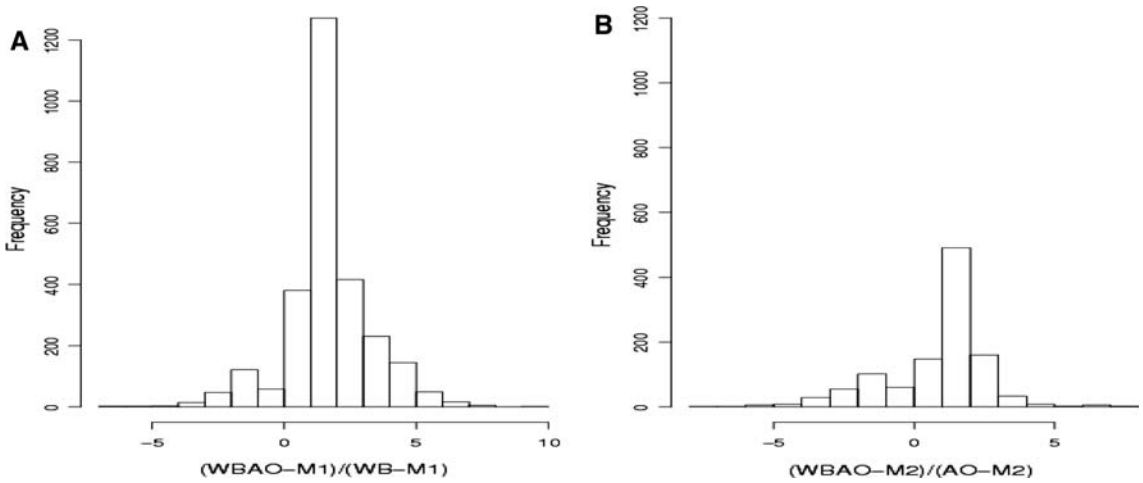

Fig. 8 Gene expression in quadruplex hybrid is more extreme than in duplex hybrids. **a** Quadruplex hybrid compared to the B73/W22 hybrid. *X*-axis: the ratio of the difference of AOWB and midparent M1 to the difference of BW and M1 (*left panel*) was divided into 15 bins of equal size; *Y*-axis: number of genes in each bin. **b** Quadruplex hybrid compared to the A188/Oh43 hybrid. *X*-axis: the ratio of the difference of AOWB and midparent M1 to the difference of AO and M1 (*left panel*) was divided into 15 bins of equal size; *Y*-axis: number of genes in each bin

parent. We note that in the quadruplex the magnitude of transgressive behavior was more extreme than in the duplex hybrid (Fig. 8). We found 2,758 genes that are transgressive both in the AOWB-M_1 and BW-M_1 comparisons. The ratio between the difference of the hybrid and the midparents (for the two types of hybrids) was used to create the bar diagram shown in Fig. 8a. For 84.1% of the genes, the transgressive behavior was more extreme in magnitude for the quadruplex (indicated by the right shift in the diagrams in Fig. 8). An identical analysis for A188/W22 found 1,105 genes to be transgressive in both hybrids. Similar to the B73/W22 analysis, there were 81.3% of genes with more extreme transgressive behavior in the quadruplex hybrids (Fig. 8b).

Discussion

Cell size increases with increasing ploidy (Rhoades and Dempsey 1966) but the stature of the homozygous materials does not change accordingly (Riddle et al. 2006), indicating that there is an increasing cell size but fewer cells in various tissues. Thus, comparable RNA expression for a particular gene indicates a proportional increase per cell per ploidy (Guo et al. 1996). The global patterns of gene expression examined here follow similar profiles as examined previously for specific genes in a ploidy series (Guo et al. 1996; Birchler and Newton 1981). In other words, there is a generalized proportional increase per cell but deviation from this trend of small magnitude occurs for many genes, which are identified in this analysis as the significant genes. The degree to which cell volume is not precisely linearly correlated with ploidy will affect the identification of such genes. The most common deviation is a greater than proportional reduction in monoploids and in ploidies above the diploid a lesser than proportional increase. This trend parallels the stature trend of the plants (Riddle et al. 2006), but we caution that this result should not necessarily be interpreted as causative. A very similar relationship of vigor in an autopolyploid series and of the global patterns of gene expression was reported for potato (Stupar et al. 2007).

In inbred versus hybrid tetraploids, patterns of gene expression show some measure of non-additivity. The magnitude of non-additive gene expression in hybrids is not great, as previously noted (Auger et al. 2005). Also as previously noted (Auger et al. 2005), the non-additive gene expression should not necessarily be interpreted as causative of heterosis. However, in the quadruplex hybrids, which typically exhibit greater heterosis than duplex hybrids for some characteristics (Riddle and Birchler 2008; Levings et al. 1967; Sockness and Dudley 1989a, b), the number and magnitude of non-additive effects increased. Heterosis involves an increase in the number of cells in a plant without dramatically altering the developmental program (East 1936), but this change is unlikely to be proportional for all cell types. Non-additive gene expression, as typically assayed, might simply reflect a different proportion of cell types in the sampled tissues from hybrids and inbreds. Alternatively, the changes could be occurring within each cell type (or a combination of both).

In conclusion, the global patterns of gene expression are on the whole consistent with previous studies of individual genes in a ploidy series (Guo et al. 1996) and inbred versus hybrid conditions (Auger et al. 2005). The spectrum of genes that are modulated from the null predictions is broad

and does not produce an obvious causative explanation for the phenotypic trends, although there are correlative trends for some genes that differ in their response in the ploidy and hybrid conditions, respectively.

References

Auger DL, Ream TS, Birchler JA (2004) A test for a metastable epigenetic component of heterosis using haploid induction in maize. Theor Appl Genet 108(6):1017–1023

Auger DL, Dogra Gray A, Ream TS, Kato A, Coe EH Jr, Birchler JA (2005) Non-additive gene expression in diploid and triploid hybrids of maize. Genetics 169:389–397

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. JRSS(B) 57:289–300

Bingham ET, Groose RW, Woodfield DR, Kidwell KK (1994) Complementary gene interactions in alfalfa are greater in autotetraploids than diploids. Crop Sci 34:823–829

Birchler JA, Newton KJ (1981) Modulation of protein levels in chromosomal dosage series of maize: the biochemical basis of aneuploid syndromes. Genetics 99:247–266

Busbice TH, Wilsie CP (1966) Inbreeding depression and heterosis in autotetraploids with application to *Medicago sativa* L. Euphytica 15:52–67

Catts VS, Catts SV, Fernandez HR, Taylor JM, Coulson EJ, Lutze-Mann LH (2005) A microarray study of post-mortem mRNA degradation in mouse brain tissue. Brain Res Mol Brain Res 138(2):164–177

East EM (1936) Heterosis. Genetics 21:375–397

Freeling M, Thomas BC (2006) Gene-balanced duplications, like tetraploidy, provide predictable drive to increase morphological complexity. Genome Res 16:805–814

Groose RW, Talbert LE, Kojis WP, Bingham ET (1989) Progressive heterosis in autotetraploid alfalfa: studies using two types of inbreds. Crop Sci 29:1173–1177

Guo M, Davis D, Birchler JA (1996) Dosage effects on gene expression in a maize ploidy series. Genetics 142:1349–1355

Guo M, Rupe MA, Yang X, Crasta O, Zinselmeier C, Smith OS, Bowen B (2006) Genome-wide transcript analysis of maize hybrids: allelic additive gene expression and yield heterosis. Theor Appl Genet 113:831–845

Hoecker N, Keller B, Muthreich N, Chollet D, Descombes P, Piepho HP, Hochholdinger F (2008) Comparison of maize (*Zea mays* L.) F1-hybrid and parental line primary too transcriptomes suggest organ-specific patterns of nonadditive gene expression and conserved expression trends. Genetics 179:1275–1283

Huang Y, Zhang L, Zhang J, Yuan D, Xu C, Li X, Zhou D, Wang S, Zhang Q (2006) Heterosis and polymorphisms of gene expression in an elite rice hybrid as revealed by a microarray analysis of 9198 unique ESTs. Plant Mol Biol 62:579–591

Kato A (1997) An improved method for chromosome counting in maize. Biotech Histochem 72(5):249–252

Kato A (1999a) Air drying method using nitrous oxide for chromosome counting in maize. Biotech Histochem 74:160–166

Kato A (1999b) Induction of bicellular pollen by trifluralin treatment and occurrence of triploids and aneuploids after fertilization in maize. Genome 42:154–157

Kato A, Birchler JA (2006) Induction of tetraploid derivatives of maize inbred lines by nitrous oxide gas treatment. J Hered 97:39–44

Levings CS, Dudley JW, Alexander DE (1967) Inbreeding and crossing in autotetraploid maize. Crop Sci 7:72–73

Mok DWS, Peloquin SJ (1975) Breeding value of 2n pollen (Diplandroids) in tetraploid × diploid crosses in potatoes. Theor Appl Genet 46:307–314

Rhoades MM, Dempsey E (1966) Induction of chromosome doubling at meiosis by the *elongate* gene in maize. Genetics 54:505–522

Riddle NC, Birchler JA (2008) Comparative analysis of heterosis in diploid and tetraploid maize. Theor Appl Genet 116:563–576

Riddle NC, Kato A, Birchler JA (2006) Genetic variation for the response to ploidy change in *Zea mays* L. Theor Appl Genet 114:101–111

Sockness BA, Dudley JW (1989a) Performance of single and double cross autotetraploid maize hybrids with different levels of inbreeding. Crop Sci 29:875–879

Sockness BA, Dudley JW (1989b) Morphology and yield of isogenic diploid and tetraploid maize inbreds and hybrids. Crop Sci 29:1029–1032

Stupar RM, Springer NM (2006) Cis-transcriptional variation in maize inbred line B73 and Mo17 leads to additive expression patterns in the F1 hybrid. Genetics 173:2199–2210

Stupar RM, Bhaskar PB, Yandell BS, Rensink WA, Hart AL, Ouyang S, Veilleux RE, Busse JS, Erhardt RJ, Buell CR, Jiang J (2007) Phenotypic and transcriptomic changes associated with potato autopolyploidazation. Genetics 176:2055–2067

Stupar RM, Gardiner JM, Oldre AG, Haun WJ, Chandler VL, Springer NM (2008) Gene expression analysis in maize inbreds and hybrids with varying levels of heterosis. BMC Plant Biol 8:33

Swanson-Wagner RA, Jia Y, DeCook R, Borsuk LA, Nettleton D, Schnable PS (2006) All possible modes of gene action are observed in a global comparison of gene expression in a maize F1 hybrid and its inbred parents. Proc Natl Acad Sci USA 103:6805–6810

Uzarowska A, Keller B, Piepho HP, Schwarz G, Ingvardsen C, Wensel G, Lubberstedt T (2007) Comparative expression profiling in meristems of inbred-hybrid triplets of maize based on morphological investigations of heterosis for plant height. Plant Mol Biol 63:21–34