Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Plant Science 174 (2008) 17–24

PLANT SCIENCE

www.elsevier.com/locate/plantsci

Gene expression non-additivity in immature ears of a heterotic F_1 maize hybrid

Giorgio Pea [a,*], Simona Ferron [a], Luca Gianfranceschi [a], Pawel Krajewski [b], M. Enrico Pè [a]

[a] Department of Biomolecular Sciences and Biotechnology, Università degli Studi di Milano, via Celoria 26, 20133 Milano, Italy
[b] Institute of Plant Genetics, Polish Academy of Sciences, Strzeszyńska 34, 60-479 Poznań, Poland

Received 1 June 2007; accepted 12 September 2007
Available online 16 September 2007

Abstract

Non-additive gene regulation has been recently suggested as an important factor promoting phenotypic variation and plasticity. In order to obtain a description of gene expression status at an early stage of ear development in a maize (*Zea mays* L.) F_1 hybrid as relative to its parental inbreds, we compared gene expression profiles in immature ears of elite inbred lines B73 and H99 to one of their F_1 hybrids (B73 × H99) using cDNA microarray technology. Results show several genes expressed at a significantly different level between both inbred lines and their hybrid. In addition, gene expression non-additivity in the hybrid was detected on a broad scale, consisting of both dominance and over-dominance components, indicating that complex non-additive interactions at the molecular level exist in the developing ear of the studied maize hybrid. Non-additively regulated genes belong to a wide range of molecular functions, indicating that several regulatory and metabolic patterns are possibly affected during ear development in the investigated hybrid. We discuss the possibility that observed gene expression non-additivity in immature ear might be an early molecular manifestation of hybrid vigor, the most exploited factor for maize agronomic improvement.
© 2007 Elsevier Ireland Ltd. All rights reserved.

Keywords: Dominance; Heterosis; Hybrid; Microarray; Non-additivity; *Zea mays*

1. Introduction

The term heterosis describes the superiority of an F_1 hybrid over its parents. Heterosis as it applies to crop breeding was first recognized by Shull in 1908 [1]. The increase of productivity that results from heterosis, or hybrid vigor, combined with the expression of adaptive traits such as fertility and resistance to biotic and abiotic stresses [2], is exploited through the development of hybrid varieties in several crop species, most markedly in maize [3]. However, the profound bases of heterosis are still elusive and the production of new hybrids still basically relies on empirical and time-consuming approaches [4]. Therefore, any added insight that could lead to the development of reliable tools for hybrid performance prediction would have an enormous impact.

Due to the complex nature of heterosis, it is generally difficult to produce reliable associations between phenotypic effects and molecular mechanisms occurring in hybrids. Therefore, the molecular bases of heterosis are still poorly understood. In fact, although quantitative genetics studies succeed in partitioning environmental and genetic effects into components of variance, their results are generally difficult to be directly associated with physiological and molecular events. In this context, a wide range quantification of intracellular molecular processes should lead to the important goal of joining quantitative genetics to genomic analysis [5,6]. Since a relevant part of biological regulations occurs at the transcriptional level, it might be possible to gain crucial information by monitoring gene expression changes on a large scale [7]. In particular, non-additive regulation in gene expression has been suggested as a potential molecular phenomenon underlying phenotypic variation in inter-specific hybrids of *Drosophila melanogaster* [8] as well as in natural hybrids and in artificial allotetraploids of *Arabidopsis thaliana* [9,10]. Therefore, it is reasonable to postulate that heterotic phenotypes in maize might also been influenced by hybrid-specific (i.e. non-additive) gene regulation. In fact, gene expression studies on maize, conducted on limited gene samples, comparing both

* Corresponding author. Tel.: +39 02 50315013; fax: +39 02 50315044.
E-mail address: giorgio.pea@unimi.it (G. Pea).

0168-9452/$ – see front matter © 2007 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.plantsci.2007.09.005

diploid versus triploid hybrids [11] and inbred lines versus their F_1 hybrids [12–15] showed that a substantial number of genes were not expressed in hybrid at the expected mid-parent value. More recently, non-additivity was also observed on a genome-wide scale in arabidopsis, maize, rice, wheat and cotton [16–23], leading to the hypothesis that gene expression variation might be related to heterosis for several traits. Two extreme modes of action for transcription regulation might be envisioned to fit with a non-additive expression model: (i) different alleles, when joined in the hybrid, are capable of combined intra-locus allelic expression; (ii) the combination of different alleles at specific regulatory loci produces genome-wide interactions leading to a general deviation in gene expression levels from the mid-parent predicted value. In maize, several data were produced supporting both the former ([24] and M. Morgante, personal communication), and the latter model [15].

We applied cDNA microarray approach to detect the presence of precocious gene expression non-additivity at an early developmental stage in ear, the organ directly involved in yield potential in maize. In particular, we set out to compare expression differences between two inbred lines (B73 and H99) and one of their F_1 hybrid (B73 $\times$ H99), which shows high level of heterosis in several vegetative and reproductive traits [25].

2. Materials and methods

2.1. Plant material

Plant material was collected from maize inbred lines B73 (Stiff Stalk Synthetic) and H99 (Illinois Synthetic 60C), as well as from their F_1 hybrid (B73 $\times$ H99). Immature upper ears were harvested from plants cultivated in open field (stage V15; http://maize.agron.iastate.edu/corngrows.html), when ear shoot tip appears and silks are just beginning to grow. For the sake of sample uniformity, only ears whose silks reached two third of the ear length were selected for all genotypes. Plant material was collected and pooled from groups of at least 10 individuals always at the same time of the day. Material was immediately frozen after removing silks and ear apexes and stored at -80 °C until RNA extraction.

2.2. Expression measurements

In this study, cDNA microarrays (print no. 606.01.04; NCBI GEO platform GPL372) produced at the University of Arizona (Tucson) were used. They contain 15606 cDNA spotted in triplicate representing 4905 *Zea mays* expressed sequence tags (ESTs) from immature ear cDNA library (strain *Oh43*; Schmidt lab, UCSD). Ten cDNA microarray slides were used, five for each tested comparisons (F_1 vs. B73 and F_1 vs. H99); according to recent literature, such a design (F_1 as common reference for contrast with both parental lines) is optimal for estimating heterosis parameters [26]. Labeling dyes were swapped in two of the five replicates for each comparison. For each genotype, RNA extracted independently from different bulks of ear tissue was used for poly(A^+) purification, retrotranscription and hybridization. Due to the described design and platforms, 15 observations were collected for each hybrid–parent comparison.

The data discussed in this publication have been deposited in NCBI Gene Expression Omnibus (GEO, http://www.ncbi.nlm.nih.gov/geo/) and are accessible through GEO Series accession number GSE2771.

2.3. Total RNA isolation and poly(A^+) purification

Tissues were ground in liquid nitrogen using mortars and pestles. Total RNA was isolated using the TRizol® protocol (Invitrogen, Carlsbad, CA), as indicated by the manufacturer (except for 5 min extra time centrifugation in TRizol® reagent), including a second step in chloroform for lowering protein contamination. For each genotype, poly(A^+) RNA was purified from 1 mg of total RNA derived from a minimum of three independent extractions using mRNA Purification Kit (Amersham Bioscience, Little Chalfont, UK). Both total and poly(A^+) RNA were tested for quality by electrophoresis on 1.5% agarose gel and quantified by absorbance at 260 nm.

2.4. Microarray hybridization

For each hybridization, 1 μg of purified poly(A^+) RNA from each genotype was independently retrotranscribed using 400 U of SuperScript II RNase H-Reverse Transcriptase (Invitrogen, Carlsbad, CA) and 2 μg Oligo$(dT)_{23}$ Anchored (Sigma–Aldrich, St. Louis, MO) as primer, in 30 μL final volume (2 h, 42 °C). cDNA probes were labeled by direct incorporation of Cy3/Cy5 modified dCTP 0.3 mM (Amersham Bioscience, Little Chalfont, UK); dATP, dGTP and dTTP 0.5 mM each, dCTP 0.2 mM. Reaction was stopped adding 1.5 μL EDTA (0.5 M, pH 8.0) and 3.75 μL NaOH (1 M) (10 min, 65 °C) and then neutralized with 0.75 μL HCl (5 M) and 9 μL Tris–HCl (1 M, pH 6.9). Probe was purified with Nucleo Spin Extract kit (Macherey-Nagel GmbH & Co., KG, Düren, Germany), protocol #4.2 with double wash in NT3 buffer. After adding 12 μg of polydeoxyadenylic acid (Amersham Bioscience, Little Chalfont, UK) the probe was lyophilized in SpeedVac™ SVC-100 H (Savant Instruments/E-C Apparatus, Holbrook, NY) and then resuspended in 29 μL Array Hyb Low Temp Hybridization Buffer (Sigma–Aldrich, St. Louis, MO) and 2 μL salmon sperm DNA (20 $\mu g\, \mu L^{-1}$). Slides were rehydrated 7 min in water-saturated atmosphere and briefly dried on heating plate (3–4 s); spotted cDNA were cross-linked to the silane-glass support applying twice 65 $mJ\, cm^{-2}$ UV light (254 nm; Stratalinker® 2400 UV cross-linker, Stratagene, La Jolla, CA). After rinsing 2 min in SDS at RT, spotted cDNA were denatured by immersion of slides in mQ water for 2 min at 95 °C. Unspecific binding sites were blocked for 40 min at 65 °C in 1% BSA, 3.5$\times$ SSC, 0.2% SDS. Slides were rinsed at RT in mQ water 50 times in each of four trays, then in isopropanol (10 immersions), and finally air-dried and stored in a clean box until hybridization. After denaturation (2 min, 98 °C in mQ water) probe was hybridized on microarray slides o.n. at 50 °C in a dark hybridization chamber (CMT-Hybridization

Chamber, Corning Inc., Corning, NY). Slides were then washed in SDS 1% solutions at increasing stringency (SSC concentration 2×, 1×, 0.1×; 10 min each, 65 °C), then in SSC 0.1× (5 min, 65 °C, twice), air-dried and stored in the dark at room temperature until image acquisition.

2.5. Signal quantification and data pre-processing

Microarray images for Cy3 and Cy5 channels were acquired using ScanArray® software on SA4000 Scanner (v3.1, Packard BioScience, Wellesley, MA), setting laser power to 90% and auto-adjusting photo-multiplicator gain to the maximum subsaturating value for each channel. Signal and background intensities and spot parameters were quantified by QuantArray® (v3.0, Packard BioScience, Wellesley, MA). Records corresponding to single bad-quality spots were manually removed. Intensity data were imported into GeneSpring (Agilent Technologies, Palo Alto, CA), where local background subtraction as well as per-chip LOWESS and 50th percentile within-array normalization functions were performed on log-two transformed data. Inter-array variability, assessed by comparing the box-plots of the intensity log ratio distributions of the replicated slides after within-array normalization [27], indicated no need for inter-slide scale normalization. Base two logarithms of hybrid/inbred expression ratios for each EST, calculated from the average over three replicate spots within each slide, were exported from GeneSpring to perform all subsequent analyses. In order to increase data reliability, we arbitrarily discarded ESTs for which less than three mean values, or less than two means for each dye-swap block, were available. As revealed by ANOVA done on intensity log ratios independently for the two comparisons, a number of ESTs showed a significant interaction with dye-swap after within-array normalization. Since the effects of incorrigible noisy data on the significance analysis and on parameter estimates were not predictable, these ESTs were also removed from gene lists before any further analysis, in order to minimize the sources of technical variation still detectable after data normalization. The described procedure also revealed a general bad quality and poor reproducibility of control spots, which were therefore all omitted from subsequent analyses.

2.6. Differential expression and heterosis parameters

Base-two logarithms of expression ratios were subjected to significance analysis in SAM v. 2.20 software [28]; one-class response analysis was applied separately to F_1-B73 and F_1-H99 data sets (5000 permutations, automatic S_0; analyses, repeated 20 times with different random seeds, indicated full reproducibility of differentially regulated ESTs lists). For each EST the estimates of additive parameter a and dominance parameter d (middle-parent heterosis) were obtained as $a = (L_2 - L_1)/2$ and $d = (L_1 + L_2)/2$ (where L_1 and L_2 are mean base-two logarithms of expression ratios of F_1 vs. B73 and F_1 vs. H99, respectively). Positive values of a indicate expression values bigger in B73 than in H99. The dominance/additivity ratio ($d/|a|$) was also calculated [29]. Evaluation of statistical significance of parameters was done by calculating standard errors of the estimates a and d as standard errors of linear functions of the means. Significance testing was done correspondingly, using an F test for linear contrasts. P-values for the families of tests corresponding to each parameter were subjected to global error analyses using a method based on fitting mixture distribution [30], allowing to estimate the false discovery rates (FDR) and false negative rates (FNR). Confidence intervals for $d/|a|$ ratios were obtained by Fieller's method [31], allowing to classify the genes into different dominance type classes. Elements of the statistical analysis involving estimation and testing linear contrasts, confidence limits calculation and mixture fitting were done by GenStat 8 (Lawes Agricultural Trust 2005, Harpenden, UK).

2.7. Real-time PCR

Primer sets for real-time PCR (optimum length 20 bp; T_m 60 °C; GC% ≥55; Table 1) were designed to the sequences of each of the tested ESTs available in GenBank, using Primer3 software [32]. Amplification products of 150–200 bp as close as possible to the 3′-end direction of coding sequence were chosen. EST sequences were checked for the presence of secondary structures using Mfold software [33] with an energy cut-off of $\Delta G_0 > -6$ kcal mol^{-1}. Total RNA was treated with deoxyribonuclease I (DNase I Amplification Grade, Sigma–Aldrich, St. Louis, MO) as reported by manufacturer, dried by vacuum speed centrifugation (SpeedVac™ SVC-100 H, Savant Instruments/E-C Apparatus, Holbrook, NY) and resuspended in 20 μl DEPC water. As internal control, primers 18S_for (5′-GACGGGT-GACGGAGAATTA-3′) and 18S_rev (5′-GCGCCCGGTATTG-TTATTTA-3′), designed on *Zea mays* 18S small subunit ribosomal RNA gene complete sequence (GenBank accession no. AF168884) were used. All cDNA syntheses were carried out by iScript cDNA Synthesis kit (Bio-Rad Laboratories, Hercules, CA) as reported by manufacturer, except for 40 min reaction time. Samples were finally diluted to 100 ng μL^{-1}. Each real-time RT-PCR reaction was carried out on 380 ng of cDNA, using

Table 1
List of real time RT-PCR primer sets

Class[a]	GenBank accession	Forward Primer (5′-3′)	Reverse Primer (5′-3′)
H99 < F_1	AI737836	ttcgctgtctcaacagcttc	tggtaggatatgcctcgaca
	AI739778	gaacccagtcgctcacttgt	tgtgtgtacagctcgcatca
	AI714900	aaccccaaccaatcacca	ccgtgaggatccctcttgt
H99 > F_1	AI737795	gggacactcatcaccacaga	catcgtgctctggaagtgg
	AI881221	ttgaaggcatagccacctgt	gtgttcaccgaagagcgaag
	AI881226	catcgtgctctggaagtgg	gggacactcatcaccacaga
B73 < F_1	AI666083	acaagcgggtagacttgctg	agctagacgagtgcgaggag
	AI691512	cacaacactggtggagttgg	aggatgtgaatgctgctgtg
	AI770730	ccgttcccactaccatacca	cagccatcttcgacattcct
B73 > F_1	AI770829	actcagaggcacttgcttgg	tcatttctcccaggtgtcg
	AI734328	ggcatcttccacttcatgct	gacgactacttcgccaccac
	AI770853	cacagggtcaaatcctcctc	cctcgtaccagttgaagtcca

[a] Class of expression as determined by SAM one-class significance analysis of microarray data (5% FDR, see Supplemental Tables 2 and 3).

the iQ SYBR Green 2× Supermix Kit (Bio-Rad Laboratories, Hercules, CA) in a 25 μL total volume. Real time RT-PCR were performed on a SBI002.0 iCycle thermo-cycler (Bio-Rad Laboratories, Hercules, CA). Amplification cycles: 3 min, 95 °C; 46 cycles 30 s 95 °C, 40 s 60 °C, 40 s 72 °C; 10 min 72 °C. Melting curve: from 55 to 95 °C, +0.5 °C increment at each cycle. For each tested EST, relative expression levels were measured from two biological replicates (each employing three repeated measurements) carried out independently on new tissue samples. Titration curves were built on the signal relative to 18S small subunit ribosomal RNA gene, employing four serial dilution of the cDNA template (1:1, 1:10, 1:100 and 1:1000). Significance testing of real-time RT-PCR results was performed by *t*-tests carried out on biological replicates averaged over technical replicates.

2.8. EST annotation and over representation analysis

Tentative contigs (TC) for the ESTs spotted on the cDNA arrays were retrieved from Maize TC Annotator whenever available (TIGR Maize Gene Index, release 16.0; http://www.tigr.org/tdb/tgi/zmgi/). Related functional and structural descriptions were also retrieved as gene ontology terms at TIGR Maize Gene Index, and updated to the last GO term definitions database at Gene Ontology Consortium (http://www.geneontology.org/, monthly release March 2007). Differentially expressed gene lists were submitted to GOSSIP v1.4.1 software [34] to test for the presence of over-represented GO categories among regulated genes (over-representation analysis or ORA). FDR <0.1 was set as significance threshold.

3. Results

3.1. Data pre-processing

Out of 4905 sample ESTs on microarray slides, 3761 (76.7%) for B73 versus F_1 and 3594 (73.3%) for H99 versus F_1 passed through quality filters. Only the 2791 ESTs (56.9%) that were shared between the two lists were considered for subsequent analyses. Correlation analysis for these ESTs showed average correlation coefficients among replicates of 0.59 (±0.09) for B73 versus F_1 and 0.32 (±0.14) for H99 versus F_1 (data not shown), which are within expected range for standard cDNA microarray experiments carried on different lines in presence of dye-labeling inversion [35]. Normalized ratio values were slightly more variable in B73 versus F_1 (average standard deviation 0.173) than in H99 versus F_1 (average standard deviation 0.103). See Supplemental Table 1 for normalized ratio values and replicate quality filtering results.

3.2. Differential expression analysis

Genes up and down regulated in the hybrid in comparison to each parental inbred line were identified by a one-class response analysis in SAM (Table 2, Supplemental Tables 2 and 3). Since the actual proportion of false positives was likely to be higher than what estimated (loss of protection due to the reduced number of hypotheses being tested because of bad-quality EST removal according to data quality check), a stringent 5% FDR significance cut-off was set. Normalized ratio values seldom indicated a two-fold or bigger expression change. Fold-change cut-off values indicating significance were 1.039/0.957 for B73 versus F_1 (*q*-value 0.0473) and 1.082/0.929 for H99 versus F_1 (*q*-value 0.0412). A higher number of differentially expressed ESTs was called out in B73 versus F_1, namely 1545 ESTs (55.4% of tested sequences) versus 394 ESTs detected in H99 versus F_1 (14.1% of tested sequences). The proportion between up and down regulated ESTs was almost balanced for B73 versus F_1 contrast, whereas it was favorable to down-regulation in hybrid for H99 versus F_1 (up to down regulated ratio 1:1.8, Table 2). Comparison between the two lists of significant genes is illustrated in Fig. 1. It shows that 165 ESTs (86 up regulated and 79 down regulated, 5.9% of total) share the same type of regulation in both inbreds; 69 ESTs (2.5% of total) show opposite regulation as to the F_1. Therefore, about 40% of ESTs regulated in H99 are H99-F_1 specific (5.8% of total), against 85% of ESTs regulated in B73-F_1 (46.9% of total).

3.3. Validation by real-time RT-PCR

In order to verify the general reliability of datasets obtained by statistical analysis of microarray data, relative expression levels of a subset of significant ESTs for each direct comparison were also determined by real-time RT-PCR performed on tissue samples collected independently from those used for microarray hybridizations. Three ESTs assigned to each class of

Table 2
Synopsis of significance analysis in parents vs. hybrid contrasts

Class[a]	Significant genes				Error rates			
	Number	Relative (%)	Total	(%)	False positives	FDR (%)[b]	False negatives	FNR (%)[c]
F_1 > B73	782	50.6	1545	(55.4)	73	4.73	317	25.4
F_1 < B73	763	49.4						
F_1 > H99	140	35.5	394	(14.1)	16	4.12	752	31.4
F_1 < H99	254	64.5						

[a] Classes of expression as determined by SAM significance analysis of microarray data (Supplemental Tables 2 and 3).
[b] Number of false positives among the significant ESTs.
[c] Number false negatives among the non-significant ESTs.

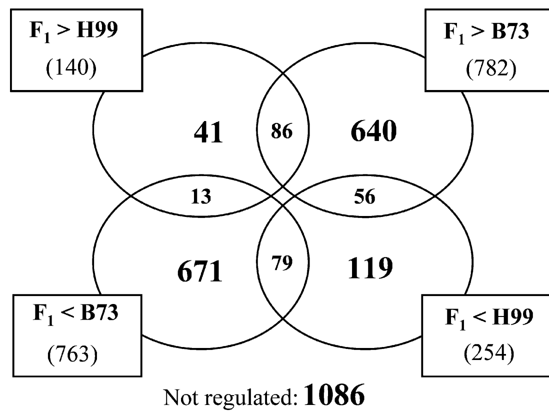

Fig. 1. Hybrid vs. parents classes of expression. Summary of classes of expression as determined by comparing significance analysis outcomes from both direct inbred vs. hybrid contrasts (SAM one-class response, 5% FDR).

regulation in direct comparisons (H99 $<$ F_1, H99 $>$ F_1, B73 $<$ F_1 and B73 $>$ F_1) were selected according to Section 2. Mean expression ratios for 10 out of 12 real-time RT-PCR comparisons were in agreement with microarray results, although in four cases they were not confirmed by significance testing ($p \leq 0.05$; Table 3). Only one EST did not significantly match microarray trend.

3.4. Additivity and dominance analysis

In order to obtain a simple description of the relationships among all three compared genotypes in terms of gene expression, additivity and dominance parameters (i.e. difference in expression levels between parental lines and between hybrid and mid-parent value, respectively), estimated from microarray data, were tested for significance as described in Section 2. A summary of results on significance ($\alpha \leq 0.05$) and sign of additive and dominant effects is reported in Fig. 2. A significant additive effect was found for 878 ESTs (FDR rate 0.076, FNR rate 0.35). A significant dominance effect was found for 1122 ESTs (FDR rate 0.070, FNR rate 0.28). Comparison of the two lists of significant genes showed that

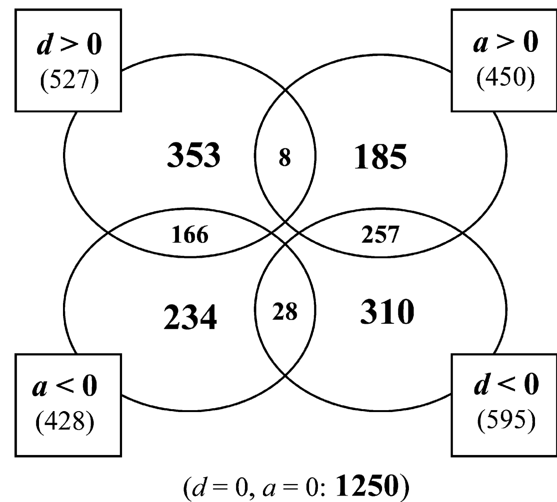

Fig. 2. Distribution of ESTs among classes defined by significance and sign of dominance (d) and additivity (a) parameters.

419 ESTs have additive effects only (15.0% of total, 47.7% of all additive), 663 ESTs have only dominant effects (23.8% of total, 59.1% of all dominant), and 459 ESTs show both dominant and additive effects (16.4% of total). Finally, 1250 ESTs (44.8% of total) are neither affected by dominant nor additive effects.

The relation between additive and dominant effects was analyzed separately for ESTs with expression similar for the two parents ($a = 0$) and the ones whose expression was different ($a \neq 0$). In the first group, 310 and 353 ESTs showed significantly negative and positive dominance respectively (Fig. 2, $\alpha = 0.05$, FDR 0.07, FNR 0.28). Thus, no preference as to the direction of dominance was observed. In the second group, ESTs were classified according to the meaning of the estimated $d/|a|$ ratio assessed by the computed confidence limits (Table 4). A vast majority (383 out of 417, 91.8%) of ESTs showing complete dominance ($d = \pm 1$) have expression levels closer to that of H99. In particular, this is true for all 31 over-dominant ESTs. The number of ESTs showing partial dominance was quite low (11). The plot of $d/|a|$ versus a for

Table 3
Microarray data validation with real-time RT-PCR

Class[a]	GenBank accession	Microarray (inbred/hybrid ratio)	Real time RT-PCR (inbred/hybrid ratio[b])
B73 $>$ F_1	AI770829	1.267	1.806**
	AI734328	1.289	1.332*
	AI770853	1.164	1.849 (ns)
B73 $<$ F_1	AI770730	0.703	1.856**
	AI691512	0.661	0.966 (ns)
	AI666083	0.546	0.011*
H99 $>$ F_1	AI737795	1.375	1.954**
	AI881221	1.303	2.378*
	AI881226	1.412	2.928*
H99 $<$ F_1	AI737836	0.763	0.621 (ns)
	AI739778	0.766	0.822 (ns)
	AI714900	0.770	1.095 (ns)

[a] Classes of expression defined according to SAM one-class statistical analysis of microarray data.
[b] One-tail t-test on biological replicate ratio averages (*$p < 0.05$; **$p < 0.01$; ns: not significant).

Table 4
Numbers of ESTs belonging to classes defined by the estimated $d/|a|$ ratio

Sign of the ratio	Type of dominance[a]	$a > 0$ (B73 > H99)	$a < 0$ (B73 < H99)	Total
Negative	Over-dominance	10	0	10
	Dominance	239	26	265
	Partial dominance	8	2	10
Zero	No dominance	185	234	419
Positive	Partial dominance	0	1	1
	Dominance	8	144	152
	Over-dominance	0	21	21
Total		450	428	878

[a] See Supplemental Table 4 for values and confidence intervals of $d/|a|$ parameter. Negative over-dominance: confidence interval for $d/|a|$ to the left of -1; negative dominance: -1 belongs to the confidence interval; negative partial dominance: confidence interval within the (-1, 0) limits; no dominance: 0 belongs to the confidence interval (positive dominance classes defined accordingly with respect to +1 value).

genes with $a \neq 0$ highlights the relationship between dominance and significant additive effects (Fig. 3). B73 and H99 are respectively the high-expressing parent for 265 (57.7%) and 194 (42.3%) of dominant ESTs. A prevalence of negative dominance for genes with positive additive effects, and vice versa, is evident. For a complete list of all determined parameters, see Supplemental Table 4.

3.5. EST annotation and enrichment of gene ontology classes

Annotations of all sample ESTs on 606.01.04 microarray slide, obtained as described in Section 2, are reported in Supplemental Table 5. Of the 2791 ESTs that passed through quality filter, corresponding to 1503 tentative contigs and 111 singletons, only 1736 ESTs (62.2%, corresponding to 810 TCs and 4 singletons) were associated to one or more GO terms (783 TCs/singletons with molecular function, 627 with biological process and 623 with cellular component; Supplemental Table 6). Annotation revealed that a wide variety of functional and structural categories were involved both in additive and non-additive gene regulation (Supplemental Table 4). To gain

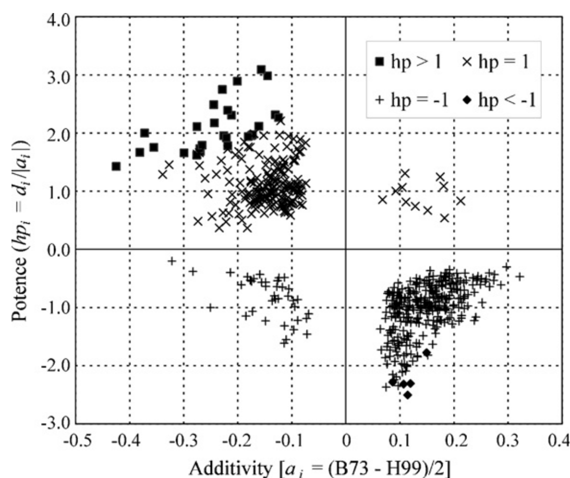

Fig. 3. Scatter plot of additivity (a) vs. $d/|a|$ ratio. Only ESTs for which $a \neq 0$ are reported. Classes defined by the estimated $d/|a|$ ratio confidence intervals are indicated by symbol code.

further insight into non-additive gene expression regulation, we tested whether certain functional patterns were over-represented among non-additively regulated ESTs by means of over-representation analysis on ESTs with $d < 0$ and $d > 0$ separately (Supplemental Table 7). Microtubule-related transport, cytoskeleton organization, GTP-binding and ion homeostasis categories were among over-represented categories within the negative dominant EST list. More GO categories resulted over-represented for positive dominant ESTs, among which organogenesis, cell cycle, response to heat stress, DNA damage repair, regulation of transcription, carbohydrate and alcohol catabolism.

4. Discussion

Several studies reported differential gene expression between inbred lines and their corresponding hybrid in maize [12–16,18,19], as well as in other species [8–10,17,22,36]. It has been proposed that transcriptional regulation might affect heterosis via either intra-locus differential allelic expression or inter-loci interactions, both leading to a general deviation in gene expression levels from the mid-parent predicted value [7]. We conducted a survey of gene expression regulation on a large sample of genes in order to determine whether non-additive gene expression variation was present between two elite maize inbred lines (B73 and H99) and their F_1 hybrid in immature ears, and to ascertain whether specific regulation of metabolic and regulatory patterns was possibly involved.

When considering complex trait variation it is reasonable to assume that even slight differences could play a major biological role [37]. Therefore, we employed a statistical threshold approach to microarray data analysis. Furthermore, in order to improve reliability of results, stringent quality controls were applied. Balancing of power and protection of analysis [38] was achieved by applying a false discovery rate (FDR) approach to determine statistic-based cut-offs [39], while keeping also under control the number of false negative.

Morphological and developmental differences are quite relevant for the considered genotypes ([25] and http://www.maizegdb.org). Therefore, immature ears were purposely collected from each genotype when they showed the same silk

length relative to the whole ear (rather than when they were of the same absolute length), so that the observed differences in gene expression between samples could be related to differences between genotypes rather than between developmental stages. In particular, immature upper ears were harvested at stage V15 (approximately 10–12 days before silk emergence), which corresponds to the beginning of the most crucial period of plant development in terms of seed yield determination (http://maize.agron.iastate.edu/corngrows.html). Both in F_1 versus H99 and F_1 versus B73 comparison, small differences in relative gene expression levels were observed, as expected for tissues at the same developmental stage. Nevertheless, a number of differentially regulated genes between parental lines and between parental lines and their hybrid were detected. Ten out of 12 ESTs tested by real time RT-PCR confirmed the trend of expression as determined by microarray data (with a proportion of comparisons confirmed by significance testing comparable to what previously observed in similar control experiments [40]).

According to recently published data [38,41], it is possible to ascertain the overall difference between compared samples (independently from the adopted statistical threshold) by estimating the total number of true positive genes (TTP) after the significance analysis as $\mathrm{TTP} = \mathrm{SL} \times (1 - \mathrm{FDR}) + \mathrm{FN}$ (where SL is the total number of genes on a significant list, FDR is the applied false discovery rate cutoff and FN is the corresponding number of false negatives). Estimated TTP values, 1130 and 1789 ESTs for F_1 versus H99 and F_1 versus B73 contrast respectively (i.e., in the order, 40% and 64% of analyzed ESTs), indicate that H99 is more similar to the F_1 hybrid than B73 at the gene expression level. Similarly, data from a recent quantitative analysis on a B73 × H99 North Carolina III design [25], indicate that female flowering time is significantly different between B73 (77 days after sowing) with respect to both H99 and the F_1 hybrid (68 and 70 days, respectively).

Microarray analyses of gene expression non-additivity in maize, conducted on various tissue samples from different inbred-hybrid systems, were recently published [16,18,19]. As in previous analysis, we detected all possible modes of gene action. Our data indicate that sign of dominance is not independent from that of additivity (*chi-squared* test, data not shown) when the latter is significant (i.e. $a \neq 0$), confirming that a parent-of-origin specificity of regulation might exist in the hybrid for alleles differently expressed between parental inbreds at the same stage [16]. Surprisingly, however, nearly 60% of significantly dominant genes (i.e. $d \neq 0$) are not differentially expressed between parental lines (i.e. $d \neq 0$ when $a = 0$). This suggests that gene expression at stage V15 of maize hybrid developing ear for the most part might be independent from that of single alleles in parental inbreds, possibly contributing to functionally relevant gene expression variability during flower development of maize hybrids.

The wide range of structural and functional roles associated with non-additive ESTs confirms that transcriptional non-additivity is widespread in the analyzed maize hybrid. At present, the task of assigning a functional meaning to these observations would be merely speculative, since a large part of analyzed ESTs still lacks any functional characterizations. However, over-representation analysis of current microarray data produced some indications that functional and structural categories might exist that are particularly enriched of ESTs showing non-additive transcriptional regulation. These might represent specific mechanisms that, being potentially associated to hybrid vigor establishment, might deserve further testing: it might be of particular interest to verify, for instance, the existence of a significant correlation between expression variability and performance in maize hybrids showing different levels of heterosis. An appealing example is set by HSP90-like chaperons (over-represented among positive dominant genes) since they were also previously reported having a central buffering role for the release of genetic variation both in *Drosophila melanogaster* and *Arabidopsis thaliana* [42–46].

We further confirmed that, as recently postulated [47], dominance and over-dominance models are not mutually exclusive at the molecular level. Our data befit the historical controversy upon genetic interpretations of hybrid vigor and corroborate the elusiveness of the genetic and molecular mechanisms most intimately involved in determining heterosis. Our work also broadens the knowledge upon gene expression in maize hybrids by unveiling novel modes of gene regulation during a crucial stage of ear development. As for previous studies, the question remains whether the observed mechanisms might participate to the determination of the heterotic phenotype. However, in perspective, this work provides pinpoints to structures and functions that, further investigated, might shed light on the molecular bases of hybrid vigor.

Acknowledgments

We wish to thank Dr. S. Berri for helping with statistical analysis issues and for critical discussions, as well as Dr. L. Mizzi for helping with software handling. This work was sponsored by Ministero dell'Università e della Ricerca Scientifica (MIUR), Cofin 2001 entitled "An Integrated Approach For Addressing Heterosis" and Cofin 2003 entitled "A Molecular and Quantitative Analysis of Heterosis in Maize".

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.plantsci.2007.09.005.

References

[1] G.H. Shull, The composition of a field of maize, Rep. Am. Breeders' Ass. 4 (1908) 296–301.
[2] T. Dobzhansky, Genetics of natural populations. XIX. Origin of heterosis through natural selection in populations of *Drosophila pseudoobscura*, Genetics 35 (1950) 288–302.
[3] D.N. Duvick, Biotechnology in the 1930s: the development of hybrid maize, Nat. Rev. Genet. 2 (2001) 69–74.
[4] J.G. Coors, S. Pandey, The genetics and exploitation of heterosis in crops, in: J.G. Coors, S. Pandey (Eds.), ASA, CSSA and SSA, Madison, WI, USA, 1999.

[5] M.K. Kerr, G.A. Churchill, Statistical design and the analysis of gene expression microarray data, Genet. Res. 77 (2001) 123–128.
[6] R.D. Wolfinger, G. Gibson, E.D. Wolfinger, L. Bennett, H. Hamadeh, P. Bushel, C. Afshari, R.S. Paules, Assessing gene significance from cDNA microarray expression data via mixed models, J. Comput. Biol. 8 (2001) 625–637.
[7] J.A. Birchler, D.L. Auger, N.C. Riddle, In search of the molecular basis of heterosis, Plant Cell 15 (2003) 2236–2239.
[8] G. Gibson, R. Riley-Berger, L. Harshman, A. Kopp, S. Vacha, S. Nuzhdin, M. Wayne, Extensive sex-specific nonadditivity of gene expression in *Drosophila melanogaster*, Genetics 167 (2004) 1791–1799.
[9] M. Vuylsteke, F. van Eeuwijk, P. Van Hummelen, M. Kuiper, M. Zabeau, Genetic analysis of variation in gene expression in *Arabidopsis thaliana*, Genetics 171 (2005) 1267–1275.
[10] J. Wang, L. Tian, H.S. Lee, N.E. Wei, H. Jiang, B. Watson, A. Madlung, T.C. Osborn, R.W. Doerge, L. Comai, Z.J. Chen, Genomewide nonadditive gene regulation in *Arabidopsis* allotetraploids, Genetics 172 (2006) 507–517.
[11] D.L. Auger, A.D. Gray, T.S. Ream, A. Kato, E.H. Coe Jr., J.A. Birchler, Nonadditive gene expression in diploid and triploid hybrids of maize, Genetics 169 (2005) 389–397.
[12] A. Leonardi, C. Damerval, Y. Herbert, A. Gallais, D. de Vienne, Association of protein amount polymorphisms (PAP) among maize lines with performances of their hybrids, Theor. Appl. Genet. 82 (1991) 552–560.
[13] T.C. Osborn, J.C. Pires, J.A. Birchler, D.L. Auger, Z.J. Chen, H.S. Lee, L. Comai, A. Madlung, R.W. Doerge, V. Colot, R.A. Martienssen, Understanding mechanisms of novel gene expression in polyploids, Trends Genet. 19 (2003) 141–147.
[14] S. Romagnoli, M. Maddaloni, C. Livini, M. Motto, Relationship between gene expression and hybrid vigor in primary root tips of young maize (*Zea mays* L.) plantlets, Theor. Appl. Genet. 80 (1990) 769–775.
[15] R. Song, J. Messing, Gene expression of a gene family in maize based on noncollinear haplotypes, Proc. Natl. Acad. Sci. U.S.A. 100 (2003) 9055–9060.
[16] M. Guo, M.A. Rupe, X. Yang, O. Crasta, C. Zinselmeier, O.S. Smith, B. Bowen, Genome-wide transcript analysis of maize hybrids: allelic additive gene expression and yield heterosis, Theor. Appl. Genet. 113 (2006) 831–845.
[17] Y. Huang, L. Zhang, J. Zhang, D. Yuan, C. Xu, X. Li, D. Zhou, S. Wang, Q. Zhang, 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 (2006) 579–591.
[18] R.M. Stupar, N.M. Springer, Cis-transcriptional variation in maize inbred lines B73 and Mo17 leads to additive expression patterns in the F1 hybrid, Genetics 173 (2006) 2199–2210.
[19] R.A. Swanson-Wagner, Y. Jia, R. DeCook, L.A. Borsuk, D. Nettleton, P.S. Schnable, 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. U.S.A. 103 (2006) 6805–6810.
[20] N.H. Syed, Z.J. Chen, Molecular marker genotypes, heterozygosity and genetic interactions explain heterosis in *Arabidopsis thaliana*, Heredity 94 (2005) 295–304.
[21] A. Uzarowska, B. Keller, H.P. Piepho, G. Schwarz, C. Ingvardsen, G. Wenzel, T. Lubberstedt, Comparative expression profiling in meristems of inbred-hybrid triplets of maize based on morphological investigations of heterosis for plant height, Plant Mol. Biol. 63 (2006) 21–34.
[22] Z. Wang, Z. Ni, H. Wu, X. Nie, Q. Sun, Heterosis in root development and differential gene expression between hybrids and their parental inbreds in wheat (*Triticum aestivum* L.), Theor. Appl. Genet. 113 (2006) 1283–1294.
[23] C.Z. Xing, Y.L. Zhao, S.X. Yu, X.L. Zhang, L.P. Guo, H.L. Wang, Relationship between gene differential expression of leaves in full opening flower stages of hybrids and their parents and heterosis in pest-resistant cotton, Yi Chuan Xue Bao 33 (2006) 948–956.
[24] M. Guo, M.A. Rupe, C. Zinselmeier, J. Habben, B.A. Bowen, O.S. Smith, Allelic variation of gene expression in maize hybrids, Plant Cell 16 (2004) 1707–1716.
[25] E. Frascaroli, M.A. Canè, P. Landi, G. Pea, L. Gianfranceschi, M. Villa, M. Morgante, M.E. Pè, Classical genetic and quantitative trait loci analyses of heterosis in a maize hybrid between two elite inbred lines, Genetics 176 (2007) 625–644.
[26] H.-P. Piepho, Optimal allocation in designs for assessing heterosis from cDNA gene expression data, Genetics 171 (2005) 359–364.
[27] Y.H. Yang, S. Dudoit, P. Luu, D.M. Lin, V. Peng, J. Ngai, T.P. Speed, Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation, Nucleic Acids Res. 30 (2002) e15–e24.
[28] V.G. Tusher, R. Tibshirani, G. Chu, Significance analysis of microarrays applied to the ionizing radiation response, Proc. Natl. Acad. Sci. U.S.A. 98 (2001) 5116–5121.
[29] D.S. Falconer, Introduction to Quantitative Genetics, Longman Scientific & Technical, London, 1989.
[30] D.B. Allison, G.L. Gadbury, M. Heo, J.R. Fernandez, C.-K. Lee, T.A. Prolla, R. Weindruch, A mixture model approach for the analysis of microarray gene expression data, Comput. Stat. Data Anal. 39 (2002) 1–20.
[31] H.P. Piepho, K. Emrich, Simultaneous confidence intervals for two estimable functions and their ratio under a linear model, Am. Stat. 59 (2005) 292–300.
[32] S. Rozen, H. Skaletsky, Primer3 on the WWW for general users and for biologist programmers, Meth. Mol. Biol. 132 (2000) 365–386.
[33] M. Zuker, Mfold web server for nucleic acid folding and hybridization prediction, Nucleic Acids Res. 31 (2003) 3406–3415.
[34] N. Bluthgen, K. Brand, B. Cajavec, M. Swat, H. Herzel, D. Beule, Biological profiling of gene groups utilizing Gene Ontology, in Genome informatics series, Proceedings of the 16th Workshop on Genome Informatics, Yokohama, Japan, 2005.
[35] G.A. Churchill, Fundamentals of experimental design for cDNA microarrays, Nat. Genet. 32 (2002) 490–495.
[36] Q. Sun, L. Wu, Z. Ni, F. Meng, Z. Wang, Z. Lin, Differential gene expression patterns in leaves between hybrids and their parental inbreds are correlated with heterosis in a wheat diallel cross, Plant Sci. 166 (2004) 651–657.
[37] W. Jin, R.M. Riley, R.D. Wolfinger, K.P. White, G. Passador-Gurgel, G. Gibson, The contributions of sex, genotype and age to transcriptional variance in *Drosophila melanogaster*, Nat. Genet. 29 (2001) 389–395.
[38] A.W. Norris, C.R. Kahn, Analysis of gene expression in pathophysiological states: balancing false discovery and false negative rates, Proc. Natl. Acad. Sci. U.S.A. 103 (2006) 649–653.
[39] Y. Benjamini, Y. Hochberg, Controlling the false discovery rate: a practical and powerful approach to multiple testing, J. R. Stat. Soc. Ser. B Stat. Meth. 57 (1995) 289–300.
[40] S.-T. Chiu, F.-J. Hsieh, S.-W. Chen, C.-L. Chen, H.-F. Shu, H. Li, Clinicopathologic correlation of up-regulated genes identified using cDNA microarray and real-time reverse transcription-PCR in human colorectal cancer, Cancer Epidemiol. Biomark. Prev. 14 (2005) 437–443.
[41] M. Langaas, B.H. Lindquist, E. Ferkingstad, Estimating the proportion of true null hypotheses, with application to DNA microarray data, J. R. Stat. Soc. Ser. B Stat. Meth. 67 (2005) 555–572.
[42] C. Queitsch, T.A. Sangster, S. Lindquist, Hsp90 as a capacitor of phenotypic variation, Nature 417 (2002) 618–624.
[43] S.L. Rutherford, S. Lindquist, Hsp90 as a capacitor for morphological evolution, Nature 396 (1998) 336–342.
[44] T.A. Sangster, S. Lindquist, C. Queitsch, Under cover: causes, effects and implications of Hsp90-mediated genetic capacitance, Bioessays 26 (2004) 348–362.
[45] T.A. Sangster, C. Queitsch, The HSP90 chaperone complex, an emerging force in plant development and phenotypic plasticity, Curr. Opin. Plant Biol. 8 (2005) 86–92.
[46] T.A. Sangster, C. Queitsch, S. Lindquist, Hsp90 and chromatin: where is the link? Cell Cycle 2 (2003) 166–168.
[47] S.W. Omholt, E. Plahte, L. Oyehaug, K. Xiang, Gene regulatory networks generating the phenomena of additivity, dominance and epistasis, Genetics 155 (2000) 969–980.