Theor Appl Genet (2010) 120:369–381
DOI 10.1007/s00122-009-1190-3

ORIGINAL PAPER

Heterotic patterns of sugar and amino acid components in developing maize kernels

Lilla Römisch-Margl · Gertraud Spielbauer · Andre Schützenmeister · Wilfried Schwab · Hans-Peter Piepho · Ulrich Genschel · Alfons Gierl

Received: 16 April 2009 / Accepted: 14 October 2009 / Published online: 7 November 2009

Abstract Heterosis is the superior performance of hybrids over their inbred parents. Despite its importance, little is known about the genetic and molecular basis of this phenomenon. Heterosis has been extensively exploited in plant breeding, particularly in maize (*Zea mays*, L.), and is well documented in the B73 and Mo17 maize inbred lines and their F1 hybrids. In this study, we determined the dry matter, the levels of starch and protein components and a total of 24 low-molecular weight metabolites including sugars, sugar-phosphates, and free amino acids, in developing maize kernels between 8 and 30 days post-pollination (DPP) of the hybrid B73 × Mo17 and its parental lines. The tissue specificity of amino acid and protein content was investigated between 16 and 30 DPP. Key observations include: (1) most of the significant differences in the investigated tissue types occurred between Mo17 and the other two genotypes; (2) heterosis of dry matter and metabolite content was detectable from the early phase of kernel development onwards; (3) the majority of metabolites exhibited an additive pattern. Nearly 10% of the metabolites exhibited nonadditive effects such as overdominance, underdominance, and high-parent and low-parent dominance; (4) The metabolite composition was remarkably dependent on kernel age, and this large developmental effect could possibly mask genotypic differences; (5) the metabolite profiles and the heterotic patterns are specific for endosperm and embryo. Our findings illustrate the power of metabolomics to characterize heterotic maize lines and suggest that the metabolite composition is a potential marker in the context of heterosis research.

Contribution to the special issue "Heterosis in Plants".

Communicated by B. Godshalk.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-009-1190-3) contains supplementary material, which is available to authorized users.

L. Römisch-Margl (✉) · G. Spielbauer · U. Genschel · A. Gierl
Lehrstuhl für Genetik, Wissenschaftszentrum Weihenstephan, Technische Universität München, Emil-Ramann-Str 8, 85354 Freising, Germany
e-mail: l.roemisch-margl@wzw.tum.de

A. Schützenmeister · H.-P. Piepho
FG Bioinformatik 340C, Institut für Pflanzenbau und Gruenland, Universität Hohenheim, Fruwirthstr. 23, 70593 Stuttgart, Germany

W. Schwab
Biomolekulare Lebensmitteltechnologie, Wissenschaftszentrum Weihenstephan, Technische Universität München, Hochfeldweg 1, 85354 Freising, Germany

Present Address:
G. Spielbauer
Plant Molecular and Cellular Biology Program, Horticultural Sciences Department, UF Genetics Institute, University of Florida, Fifield Hall, PO Box 110690, Gainesville, FL 32611, USA

Present Address:
U. Genschel
Bayerische Patentallianz GmbH, Destouchesstr. 68, 80796 Munich, Germany

Introduction

Heterosis results in the phenotypic superiority of a hybrid over its parents with respect to traits such as growth rate, reproductive success, and yield. Despite its importance, the genetic and the molecular basis of heterosis remains unsolved. Heterosis does not affect individuals as a whole, but, as a rule, the expression of heterosis is character specific. A hybrid may be superior to, equal to, or inferior to its parents (Brieger 1950) for a specific heterotic character under different conditions and for different inbred crosses. Heterotic patterns for the same character vary among different species. For example, heterosis has much stronger and more ubiquitous effects in maize than in *Arabidopsis* (Zanoni and Dudley 1989; Meyer et al. 2004; Springer and Stupar 2007).

Hybrid vigor was first studied systematically by Josef Gottlieb Kölreuter in the 1760s (1766) and has been recognized in many plant species during the last centuries. Darwin (1876) established two basic techniques for the development of the hybrid corn program: cross-pollination and controlled self-pollination. The rediscovery of Mendel's laws around 1900 and the cumulative effect of many experiments demonstrating the positive effects of cross-breeding, led to the heterosis concept in the early twentieth century. Heterosis was first recognized and its theory was formulated in the work with hybrid corn in 1908 and 1911 (Shull 1908, 1911, 1952). Shull's concept of heterosis states that different alleles when combined in heterozygotes exert a complementary physiological action, resulting in increased size, yield and vigor. The term heterosis, coined by Shull (1914, 1952), was enlarged by evolutionists to include heterosis in survival, i.e., adaptive, selective and reproductive advantage (Dobzhansky 1950; Tsaftaris 1995).

The basic principles of heterosis were already established in the first decades of the twentieth century, but the possible explanations of the genetic base of heterosis are still controversial. The overdominance and epistasis hypotheses suggest that no specific genes or alleles are responsible for hybrid vigor, whereas the dominance and pseudo-overdominance hypotheses suggest that individual genes having specific phenotypic effects are involved. It is important to bear in mind that these terms are not connected with molecular principles, as they were coined before molecular genetic concepts were formulated. Although many investigators favored one hypothesis over the other, the dominance, pseudo-overdominance, overdominance, and epistasis hypotheses are not mutually exclusive and heterosis could result from all four kinds of gene action (Lippman and Zamir 2007).

Furthermore, the genetic, physiological or metabolic basis of heterosis has been debated for nearly a hundred years without an emerging consensus. The difficulty in formulating the genetic basis of heterosis has at least two major causes: first, in most cases multiple genes contribute to the response of the F1 hybrid, and second, the multiple genes interact in ways that mask the action of each other. Additionally, the phenotypic performance depends on both genetic and non-genetic influences on plant development. Many quantitative traits in maize, including grain yield, are highly integrative and show significant variation attributable to genotype–environmental interactions (Griffing and Zsiros 1971; Moll et al. 1965; Rhodes et al. 1992; Stuber 1994).

Recent studies on recombinant inbred lines (RIL) and introgression lines (IL) of a cross between the *Arabidopsis thaliana* accessions Col-0 and C24 analyzed the metabolic signature to elucidate this biological phenomenon at the metabolic level and its impact to predict biomass production. Lisec et al. (2009) reasoned that metabolite heterosis may be important for biomass accumulation given that a linear combination of metabolite levels significantly correlated with biomass heterosis. They showed that the metabolite profiles of parents contain significant information about biomass heterosis displayed by the cross. Particularly, the combination of genetic markers with metabolite markers can lead to the improvement of heterosis prediction and can help to complete our understanding of the molecular basis of heterosis (Gärtner et al. 2009; Steinfath et al. 2010).

Currently, very little is known about the metabolite patterns in hybrids of agronomically important plants. Maize provides an excellent system for the study of heterosis. Lorenz (1972, 1975) demonstrated that the concentrations of certain amino acids in maize leaves correlate with hybrid yield. He found that relatively high yielding double-cross and three-way cross hybrids were predominantly obtained when single-cross hybrids and inbred lines with markedly different amino acid compositions were used as parents.

In the present study, we analyzed the metabolite patterns in developing maize kernels for a model hybrid B73 × Mo17 and its parental lines (B73, Mo17). Limited information is available regarding the heterotic patterns in maize kernels, especially in maize endosperm. The double fertilization process of flowering plants results in the diploid embryo and triploid endosperm. Because of this unique genetic composition in the endosperm, the gene dosage effect can play a major role and thus contribute to hybrid vigor (Guo et al. 2003; Birchler et al. 2007).

A genome-wide mRNA profiling of maize hybrids and their inbred parents (among others B73 and Mo17 lines) demonstrated that the overall level of gene expression in the maize endosperm was dosage dependent, e.g., the gene expression was predominantly proportional to the parental gene contribution of $2n$ maternal and $1n$ paternal, but also

dominant and overdominant expression patterns were revealed (Guo et al. 2003). Similarly, in the hybrid embryos the expression patterns were predominantly additive but also dominant and overdominant expression patterns have been found (Stupar and Springer 2006; Meyer et al. 2007b). As shown by Singletary et al. (1997), dosage effects were associated with metabolite patterns, especially with carbohydrate and starch biosynthesis in endosperm of starch-deficient mutants of maize. Genes related to metabolism appear to be associated with different heterotic patterns in the maize embryo, endosperm and during grain filling (Meyer et al. 2007b; Meng et al. 2005; Wang et al. 2007; Jahnke et al. 2010).

In our previous work, we quantified the relative flux contributions of the major pathways of central carbohydrate metabolism in the maize kernel to starch synthesis (Spielbauer et al. 2006). The carbohydrate flux pattern of maize kernels turned out to be very robust to genetic perturbation and was not associated with heterosis. Thus, relative carbohydrate fluxes have little value as markers for heterosis in the maize kernel. Therefore, in this study, we focused on metabolite pool sizes as potential markers for heterosis with the intention to answer the following questions:

1. Do the B73, Mo17, and B73 × Mo17 genotypes differ in their metabolite pattern?
2. Which heterotic patterns are detectable at the level of individual metabolites?
3. Is the metabolite pattern dependent on the developmental stage?
4. What is the difference between the metabolite patterns in embryo and endosperm tissues?

To this end, we surveyed the levels of starch and protein components and 24 metabolites such as carbohydrates and amino acids that are key intermediates for starch and protein kernel storage products in the B73 and Mo17 inbred lines and in the B73 × Mo17 hybrid. Metabolites were monitored throughout grain filling at 8–30 DPP, the most crucial period of plant development in terms of seed yield.

Materials and methods

Plant material

Maize plants of the inbred lines B73 and Mo17 were grown in the greenhouse in summer 2006. One ear per plant was hand pollinated to obtain B73, Mo17, and B73 × Mo17 genotypes. Ears were harvested consecutively at 8, 12, 16, 20, 25, and 30 DPP. The kernels from individual cobs were pooled, lyophilized, and stored at −70°C. Dry matter content and specific metabolites were quantified based on 2–5 biological replicates for each genotype, time point, and tissue type. Kernels older than 12 DPP were dissected into endosperm, embryo, and pericarp, which were analyzed separately. For kernels up to 12 DPP this dissection was not possible.

Dry matter and chemical analysis

Dry matter content was determined by weighing a subset of 10–30 kernels (8–30 DPP) before and after lyophilization. Dry weights of endosperm and embryo (16–30 DPP) were determined by weighing dissected samples from 10 kernels after lyophilization.

Starch content was determined from maize kernels according to the total starch assay procedure (Megazyme). Carbohydrates were extracted from maize kernels with 75% methanol at 4°C. An ICS-3000 system (Dionex) equipped with a CarboPac PA1 column (250 × 2 mm, Dionex) and a PA1 guard column (50 × 2 mm, Dionex) was used to separate the compounds at a flow rate of 0.25 ml/min at room temperature. For separation of glucose, fructose, and sucrose isocratic elution was carried out with 100 mM NaOH. Phosphorylated compounds were eluted using a linear NaOH/Na-acetate gradient in an isocratic background of 18 mM NaOH. The Na-acetate concentration was kept at 0 mM from 0 to 5 min, increased linearly from 0 to 500 mM between 5 and 45 min, then at 500 mM for an additional 10 min, and was then returned to the initial condition from 55 to 65 min. The detection was carried out by a pulsed amperometry cell equipped with a working gold electrode and a combined pH-Ag/AgCl reference electrode. The detector was used in the integrated amperometry mode with the preset standard quad potential-time sequence for carbohydrate detection. Quantification of glucose, sucrose, fructose, glucose 6-phosphate, fructose 1-phosphate, and fructose 6-phosphate was achieved by recording calibration curves from original metabolite standards.

Total protein extraction was performed through a modified two-step protein extraction protocol according to Cho et al. (2006). Finely ground samples from whole kernels (100 mg, 8–12 DPP), endosperms (100 mg, 16–30 DPP), or embryos (20 mg, 16–30 DPP) were pre-extracted three times with 2 ml of cold acetone at 4°C, followed by centrifugation at 3,000 rpm for 10 min at 4°C. After removal of acetone the pellets were air-dried. For solubilization of proteins the pellets were mixed with 2 ml of a buffer containing 9 M urea, 18 mM Tris–HCl pH 8.0, and 50 mM DTT. This mixture was incubated for 5 min at room temperature. The suspension was centrifuged at 3,000 rpm for 10 min and the supernatant was collected. This procedure was repeated four times and the supernatants were pooled. The combined supernatant was used for

protein quantification with a kit (Bio-Rad Laboratories Inc., USA).

Amino acids were extracted from dried plant material. A total of 50 mg of cornmeal derived from whole kernel (8–12 DPP) or from endosperm (16–30 DPP) or 20 mg of powdered embryo (16–30 DPP), respectively, were extracted twice with 1 ml of 25% acetonitrile. The sample preparation procedure was carried out according the EZ:faast kit for free (physiological) amino acid analysis (Phenomenex). The amino acids were analyzed on a Thermo Finnigan Trace DSQ mass spectrometer coupled to a 0.25 µm SLB5MS (SUPELCO) fused silica capillary column with a 30 m × 0.25 mm inner diameter. Helium (1.1 ml min^{-1}) was used as a carrier gas. The injector temperature was 220°C, set for split injection, with a split ratio of 1:10. The temperature program was 110°C for 3 min, 110–140°C at a rate of 5°C min^{-1}, 140°C for 18 min, and 140–320°C at 10°C min^{-1}. The ion source temperature was 250°C. Mass range was recorded from *m/z* 50–650 and spectra were evaluated with the Xcalibur software version 1.4 supplied with the device.

Statistical methods

Estimating of LS means, pairwise contrast, and heterosis pattern

Starch is the most abundant metabolite in the maize kernels. To transform the measured values such that they are independent of the starch content, we used following correction for an originally measured metabolite value x_{original}:

$$x_{\text{corrected}} = \frac{x_{\text{original}}}{(100 - \%\text{Starch})/100}$$

This type of correction was chosen because the starch content was originally measured as percentage of starch for the total kernel dry matter weight (DW), measured in mg (Starch)/100mg (DW). Starch is mainly stored within the endosperm of the maize kernel. Thus, this correction was carried out only for the complete kernel and for the endosperm measurements, but not for the embryo.

Before obtaining genotypic estimates and heterosis estimates, we looked at the residuals obtained from fitting a linear mixed model. Any outlying observations that influenced the resulting quantile–quantile (QQ) plot or residual plot of standardized residuals were removed. To further assess normality of residuals, we simulated data sets using the variance components as estimated for a specific metabolite, fitted the model again for each set of simulated data, obtained those residuals for the simulated data and plotted them. This allowed comparing QQ plots of the original data with or without outlying observations to what one would expect considering the underlying correlation structure of the residuals. Using several simulated data sets allows plotting an envelope around the observed residuals. This makes it easier to assess normality of these residuals.

To assess mid-parent heterosis for a given metabolite, we fitted a linear mixed model with SAS software version 9.2 separately for each metabolite. This facilitated computing three linear pairwise contrasts, one mid-parent heterosis contrast, and the least squares means (LS means) of each genotype-date combination in one step. We used a logarithmic transformation for all metabolites because the majority of metabolites exhibited some kind of variance heterogeneity, which was obvious in the plots of predicted values versus standardized residuals. To make statistical inference consistent for all metabolites we transformed each variable.

The linear mixed model we used was:

$$y_{ijklm} = \log(x_{ijklm}) = \alpha_i + \beta_j + \gamma_{ij} + a_{ik} + e_{ijklm}$$

α_i Fixed effect of the *i*th genotype, with $i \in \{1, 2, 3\}$,
β_j Fixed effect of the *j*th date, with $j \in \{1, 2, 3, 4, 5, 6\}$,
γ_{ij} Crossed effect of *i*th genotype and *j*th date,
a_{ik} Random effect of the *k*th plant of the *i*th genotype, with $a_{ik} \overset{i.i.d.}{\sim} N(0, \sigma_p^2)$,
e_{ijklm} Residual of measurement *m*, genotype *i*, date *j*, plant *k*, and sample *l*, with $e_{ijklm} \overset{i.i.d.}{\sim} N(0, \sigma_e^2)$.

Because of the sampling strategy that was applied, this design is unbalanced. We also considered fitting a random sample effect, but due to the imbalance there were too few degrees of freedom to reliably estimate the variance for this effect. The residual degrees of freedom for Wald-type *F* and *t* tests were approximated by the Kenward–Roger method (Kenward and Roger 1997), which is available in PROC MIXED of SAS 9.2 software.

To compute a mid-parent heterosis contrast of the embryo values, we used the estimates of all three genotypes and specified the null-hypothesis:

$$H_0 : AB - \frac{A}{2} - \frac{B}{2} = 0$$

To compute the average of parents' contrast of the endosperm values, we used the estimates of all three genotypes and specified the null-hypothesis:

$$H_0 : AB - \frac{2A}{3} - \frac{B}{3} = 0,$$

where *AB* represents the hybrid B73 × Mo17, *A* the inbred B73, and *B* the inbred Mo17. This contrast can be interpreted as the weighted mid-parent heterosis contrast with respect to the allelic composition of the endosperm tissue (Guo et al. 2003).

If β is the vector of fixed effects, a linear contrast, which tests the hypothesis of equality is defined as $L'\beta = 0$. The mid-parent heterosis and average of parents contrasts, as well as pairwise linear contrasts, are tested using an approximately t distribution for the test statistic:

$$t = \frac{L'\hat{\beta}}{\sqrt{L'\hat{C}_{\beta}L}}$$

where the denominator degrees of freedom were approximated using the method of Kenward and Roger (1997). $\hat{C}_{\beta}$ is the fixed-effects part of an estimate of the generalized inverse of the coefficient matrix in the mixed model equations (SAS-Doc PROC MIXED) and $\hat{\beta}$ represents the estimate of β.

We also applied all three pairwise linear contrasts among genotypes at each time. The null hypotheses all have the same form:

$$H_0 : A - B = 0,$$

where A and B differ such that each genotype was tested with the two other genotypes at each time.

The mid-parent heterosis and average of parents contrasts as well as pairwise linear contrasts obtained from fitting the mixed model for each genotype–time combination were used to classify the heterosis effects. Additivity was assumed when the F1 genotype mean was not significantly different from the mean of the parents. Overdominance and underdominance were assumed when the F1 genotype mean was significantly larger or smaller than the means/weighted means of both B73 and Mo17 genotypes. High-parent or low-parent dominance was inferred when the F1 genotype mean was not significantly different from one parent and significantly larger or smaller than the other parent. We identify further the dominance effects as maternal or paternal effect.

Classification of metabolite patterns

Two different strategies were applied and both aim at classifying and visualizing similarities as well as differences among genotype–time combinations. We used the LS means of each metabolite computed for a specific date, fitted with the mixed model shown above.

The first method that we applied was hierarchical bottom-up clustering with complete linkage and Euclidean distances (Murtagh 1985). The cluster agglomeration method *complete linkage* computes cluster distances as the distance between the most distant members of two clusters. We used the Euclidean distance because it is widely accepted and easy to interpret. To remove any bias introduced by scale differences among these metabolite measurements, we first standardized all values to zero mean and unit variance for each metabolite.

A second method of visualizing similar information inherent to the metabolite pattern of LS means estimates is using PCA plots. At first, a singular value decomposition (SVD) of the standardized data matrix Z was performed:

$$Z = USV',$$

where U is an $(n \times p)$ orthonormal matrix, S is a diagonal matrix of order p, and V' is the transpose of a $(p \times p)$ orthogonal matrix V. S contains the singular values of the decomposed matrix Z as diagonal elements, which are arranged so that $s_1 \geq s_2 \geq \cdots \geq s_p \geq 0$.

V defines a rotation of the original axes to a new set of axes within principal component (PC) space, which can be obtained by post-multiplying matrix Z by V (Digby and Kempton 1987; Gower and Hand 1996). This factorization represents Euclidean distances of the objects (genotype–time point combinations in our case) within the PC space.

Reconstruction of the original data

For developmental stages DPP16, DPP20, DPP25, and DPP30 amino acids and DW were separately measured in two different tissues of the maize. This allowed us to reconstruct the values for the whole maize kernel because the abundance of a specific amino acid was measured in nmol/mg DW. Let $\omega^{\text{en}}, \omega^{\text{my}}$ be the dry matter weights of endosperm and embryo, respectively.

If $\alpha_i^{\text{endo}}, \alpha_i^{\text{embryo}}$ represent the measurements of the ith amino acid in both tissues, the value for the whole maize kernel α_i^{ker} can be reconstructed using:

$$\alpha_i^{\text{ker}} = \frac{\omega^{\text{endo}}\alpha_i^{\text{endo}} + \omega^{\text{embryo}}\alpha_i^{\text{embryo}}}{\omega^{\text{endo}} + \omega^{\text{embryo}}}$$

To make it consistent with the starch correction, we corrected the endosperm dry matter weight before applying the reconstruction of the complete maize kernel. Each original dry matter weight $\text{DW}_{\text{original}}$ measured for the endosperm was corrected as:

$$\text{DW}_{\text{corrected}} = \text{DW}_{\text{original}} \times (100 - \%\ \text{Starch})/100.$$

For making interpretations on the original scale, the adjusted means and their corresponding standard errors were back-transformed using:

$$y = \exp x$$

and

$$\text{SE}_y = \exp x \times \text{SE}_x,$$

where x is a value on the original scale, y is the value after log-transformation, and SE is the standard error (Piepho

2009; Connolly and Wachendorf 2001; Taylor 2006). Adjusted means can be interpreted as estimates of medians and this property survives after back-transformation. The back-transformed means can therefore be explicitly labeled as estimates of medians on the original scale.

Results

High-parent heterosis in dry matter content of maize kernels

Dry matter content increased in a nearly linear fashion for the three genotypes. At all time points between 8 and 30 DPP the dry matter content of B73 × Mo17 kernels was higher than in either parental line (Fig. 1a). The differences between Mo17 and the hybrid were significant throughout the time interval studied, while the differences between B73 and the hybrid were significant at four time points. Thus, high-parent heterosis in the B73 × Mo17 system is manifest in dry matter content throughout kernel development.

Analysis of starch, sugar, protein, and free amino acid content in maize kernels

Genotypic differences

The global trends of metabolite levels over time were similar in the kernels of the three genotypes. Starch was typically the most abundant storage compound of the maize kernels. Among the metabolites analyzed, sucrose, glucose, and fructose were the main sugar components, while alanine, asparagine, aspartate, glutamine, glutamate, and proline prevailed as the most abundant free amino acids. Phosphorylated sugars and the remaining amino acids were approximately two to three orders of magnitude less abundant (Supplementary Fig. 1).

The content of the sugar and amino acid components, expressed on a dry weight basis, is strongly affected by the changes in starch levels. In order to avoid the "dilution effect" of starch accumulation in endosperm on other metabolites, their concentrations were calculated on a starch-corrected dry weight basis.

Among all metabolite–time combinations, the highest metabolite concentration determined in any genotype was observed most frequently in Mo17 (57%) and much less frequently in B73 (21%) or in B73 × Mo17 (22%). For instance, glucose and fructose are predominantly higher in Mo17 than in B73 or B73 × Mo17 during kernel development (Fig. 1). Similarly, alanine, asparagine, lysine, threonine, and tryptophan levels were constantly more abundant in Mo17 than in the other genotypes. Contrary to this bias, the highest abundance of starch was found in B73 × Mo17 kernels at nearly all time points. Other notable exceptions were sucrose and methionine with the highest concentration in B73 × Mo17 (Fig. 1 and Supplementary Fig. 1). Among the pairwise comparisons

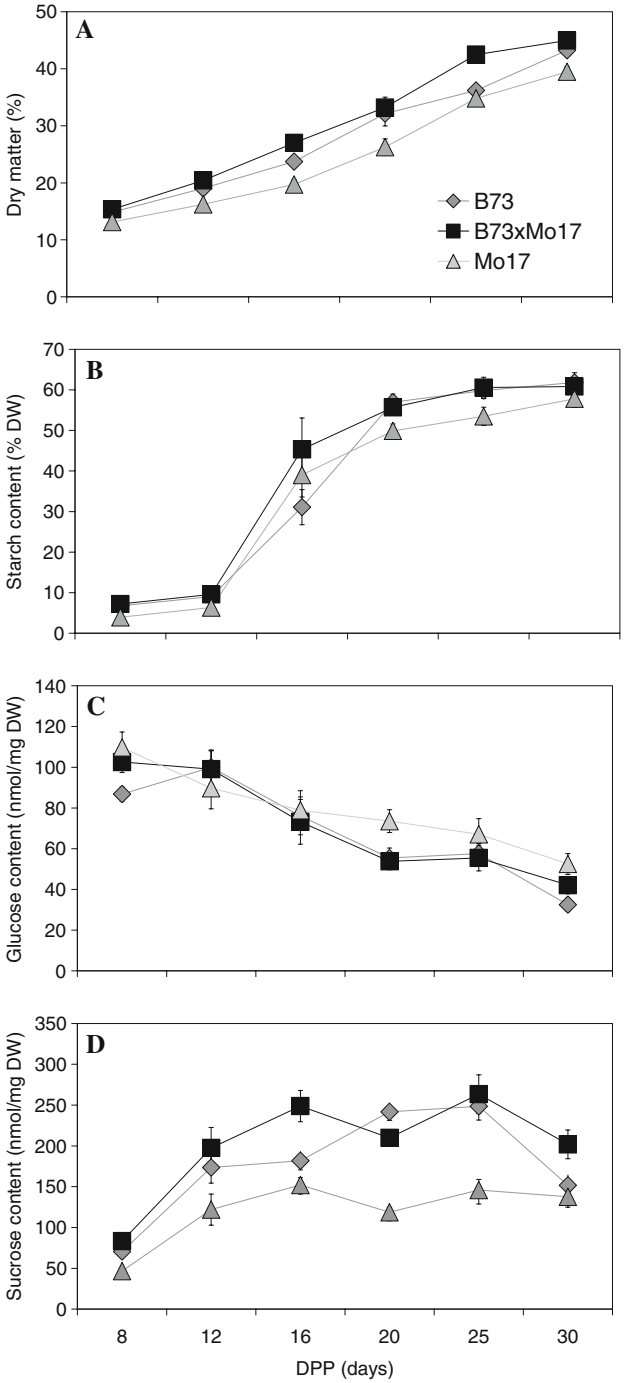

Fig. 1 Dry matter (**a**), starch (**b**) glucose (**c**) and sucrose (**d**) content in B73, B73 × Mo17 and Mo17 maize genotypes. The diagrams illustrate the reconstructed original data as described in "Material and methods". Glucose and sucrose contents are expressed on a starch-corrected dry weight basis

of genotypes (for a given metabolite and time) most of the significant differences were found between Mo17 and either B73 or B73 × Mo17. The abundance of significant differences between B73 and B73 × Mo17 was remarkably lower and their metabolite compositions are rather similar (data not shown).

The time-independent LS means characterize the level of each metabolite over the time analyzed and clearly show the overall genotypic differences (Fig. 2). Among the 26 compounds quantified, the highest time-independent concentration was frequently observed in Mo17 (16 metabolites). Out of the remaining ten metabolites the highest concentration of aspartate, glutamate, phenylalanine, proline, serine, and tyrosine was observed in B73 while the highest level of sucrose, starch, glucose-6-phosphate, and methionine was characteristic for the B73 × Mo17 genotype. The differences in the time-independent LS means between Mo17 and the two other genotypes were abundant. Significant differences were observed for 14 metabolites between Mo17 and B73 and for 11 metabolites between Mo17 and B73 × Mo17. Sucrose and lysine were significantly different among all three genotypes. Glucose, fructose, and starch as well as alanine, asparagine, and tyrosine showed significant differences between Mo17 and the other two genotypes. In contrast glucose-6-phosphate in B73 and glutamine in B73 × Mo17 had a significantly lower level compared with the other two genotypes.

Analysis of heterotic patterns

An additive mode of gene action would predict the level of a trait in the F1 hybrid to be equal to the average of the parents. In triploid endosperm, the expected level would be proportional to the parental contribution, i.e., the average of parents: Ave = $(2P_{\text{female}} + 1P_{\text{male}})/3$ (Guo et al. 2003). We took advantage of this mode of calculation and adopted it for metabolite levels in whole kernels, given that the main kernel part is the endosperm (the embryo represents only 3–10% of the kernel dry weight throughout the development).

In order to analyze heterotic patterns of individual metabolites, we calculated the average of parent values for all metabolite traits across the developmental stages in whole kernels. Approximately 10% of the individual hybrid metabolite concentrations were significantly different from the respective weighted parental means. The concentrations of these nonadditive metabolites were frequently smaller, than the calculated average of parent value.

The calculation with time-independent LS means gave a highly similar picture (Table 1). Among the seven metabolites with nonadditive pattern four amino acids: asparagine, glutamine, serine, and threonine were below and starch, sucrose, and glucose-6-phosphate were above the weighted parental means. The heterotic patterns of these

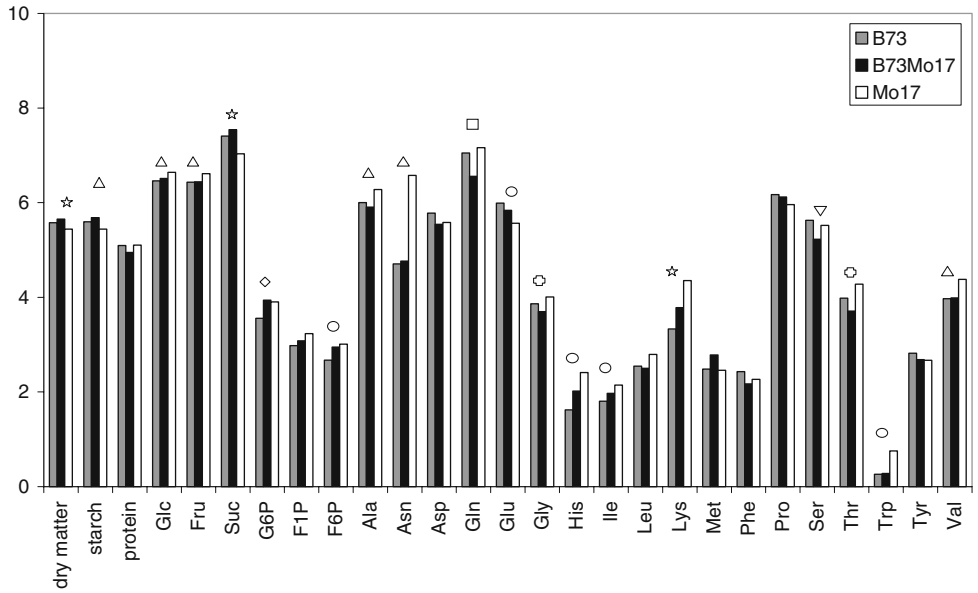

Fig. 2 Time-independent least square means of dry matter, starch and protein as well as sugar and amino acid levels in B73, B73 × Mo17 and Mo17 maize genotypes. The *bars* represent the data set in %DW for dry matter, starch and protein and in nmol/mg DW for sugar and amino acid compounds after log-transformation. The log-transformation provides the presentability of compounds with different concentrations of several magnitudes on one diagram. The original values were multiplied by 10 to avoid values below zero. Significant differences between genotypes are indicated as follows: *stars* among the three genotypes, *triangle* between Mo17 and B73 as well as Mo17 and F1 hybrid, *diamond* between B73 and F1 as well as B73 and Mo17, *square* between F1 and B73 as well as F1 and Mo17, *open circle* between B73 and Mo17, *inverted triangle* between B73 and F1, *cross* between Mo17 and F1

metabolites were further classified. Sucrose exhibited overdominance, starch and glucose-6-phosphate exhibit high-parent dominance. Asparagine, serine, and threonine exhibited low-parent dominance and glutamine exhibited clear underdominance. We identified three cases of maternal dominance for starch, asparagine, and threonine and two cases of paternal dominance for serine and glucose-6-phosphate.

Time courses of carbohydrate and amino acid components

Starch deposition accelerated rapidly after 12 DPP and leveled off at 20 DPP in all genotypes (Fig. 1b). We also observed that both the total sugar concentration and the total amino acid concentration reached a maximum at 12 DPP, i.e., at the onset of the rapid starch filling phase and then declined continuously (data not shown). This overall pattern fits the expectation that the synthesis of starch and other storage compounds depletes the endosperm of low-molecular weight precursors.

Three basic types of time courses for individual metabolite concentrations could be distinguished. The metabolites in the first group showed a maximum level at the early stage and then declined continuously towards maturity. This group includes glucose, fructose, isoleucine, serine, and valine (Fig. 1c; Supplementary Fig. 1). The levels of metabolites in the second group, like sucrose, methionine, and lysine, showed an increasing trend with age (Fig. 1d; Supplementary Fig. 1). Metabolites in the third group passed through a maximum level between 12 and 16 DPP. This group includes sugar-phosphates, alanine, leucine, glycine, histidine, threonine, and tryptophan (Supplementary Fig. 1). The distribution of the significant differences in metabolite content between genotypes or heterotic patterns showed no clear trend over development.

Tissue specificity

In order to inspect the contribution of endosperm and embryo in the metabolite pattern of the maize kernel, we analyzed the amino acid and protein content in dissected kernels. The concentrations of individual free amino acids vary over three orders of magnitude across genotypes, developmental stages, and tissues (Fig. 3). In both endosperm and embryo tissues, glutamine, glutamate, alanine, and aspartate were among the most abundant amino acids. Conspicuously, proline was by far the most abundant amino acid in embryo; its concentration was two orders of magnitude higher in the embryo than in endosperm. Thus, 90% of proline in the maize kernels at 30 DPP is concentrated in embryo and only 10% in endosperm.

As was observed in whole kernels (see above), the Mo17 genotype is predominantly associated with the highest amino acid levels in endosperm and in embryo

Table 1 Heterosis effects different from additivity

Sample type	Trait	Probt	Difference from parental main[a] (%)	Range of genotypes	Pairwise comparison			
					B73 versus B73 × Mo17	Mo17 versus B73 × Mo17	B73 versus Mo17	Dom. pattern
Whole kernel	Starch	0.0157	16.08	H > M > P	M ≈ H	P <> H	M <> P	HP/MD
Whole kernel	Sucrose	<0.0001	27.67	H > M > P	M <> H	P <> H	M <> P	O
Whole kernel	Asparagine	0.0049	−62.42	P > H > M	M ≈ H	P <> H	M <> P	LP/MD
Whole kernel	Serine	0.0060	−30.60	M > P > H	M <> H	P ≈ H	M ≈ P	LP/PD
Whole kernel	Glutamine	0.0146	−41.09	P > M > H	M <> H	P <> H	M <> P	U
Whole kernel	Threonine	0.0347	−31.79	P > M > H	M ≈ H	P <> H	M <> P	LP/MD
Whole kernel	Glucose-6-P	0.0370	28.64	H > P > M	M <> H	P ≈ H	M ≈ P	HP/PD
Endosperm	Valine	0.0072	−52.03	P > M > H	M ≈ H	P <> H	M <> P	LP/MD
Endosperm	Serine	0.0049	−62.42	P > M > H	M ≈ H	P <> H	M <> P	LP/MD
Endosperm	Asparagine	0.0060	−30.60	P > H > M	M ≈ H	P <> H	M ≈ P	LP/MD
Endosperm	Leucine	0.0146	−41.09	P > M > H	M ≈ H	P <> H	M <> P	LP/MD
Embryo	Asparagine	0.0072	−52.03	H > M > P	M ≈ H	P <> H	M <> P	LP/MD
Embryo	Methionine	0.0049	−62.42	P > H > M	M ≈ H	P <> H	M <> P	HP/MD

Probt P value of the significance test between the parental mean and the hybrid LS mean. Range of genotypes show the downward sequence of the LS means in the three genotypes: *H* hybrid, *M* B73, *P* Mo17, *O* overdominance, *U* underdominance, *HP* high-parent dominance, *LP* low-parent dominance, *MD* maternal dominance, *PD* paternal dominance, <> significantly different metabolite concentration, ≈ not significantly different metabolite concentration

[a] average of parents for kernels and endosperm and mid-parent heterosis for embryo is calculated from the reconstructed original data as described in "Material and methods"

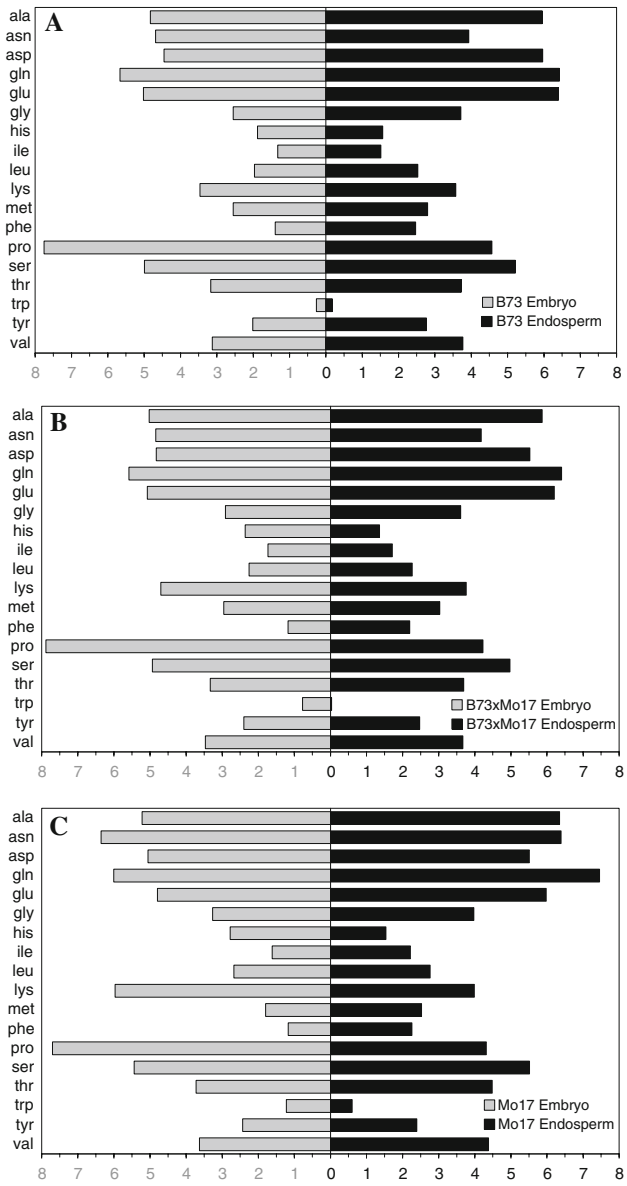

Fig. 3 Comparison of free amino acid content in embryo and endosperm of maize kernels in B73 (**a**), B73 × Mo17 (**b**) and Mo17 (**c**) genotypes using the log-transformed time-independent LS means as described in "Material and methods". The *bars* represent the data set in nmol/mg DW for sugar and amino acid compounds after log-transformation. The original values were multiplied by 10 to avoid values below zero. The scale of the *X*-axis is mirrored for embryo (*grey label*) and for endosperm (*black label*)

(Supplementary Figs. 2, 3). In both tissues the most significant differences occurred between Mo17 and B73 as well as between Mo17 and B73 × Mo17. With respect to the genetic composition in the endosperm and in the embryo the average of parents and mid-parent heterosis were calculated, respectively. Table 1 shows that additivity remained the most prevalent heterotic pattern in both of the dissected kernel parts. In endosperm, we found four amino acids representing low-parent dominance (valine, serine, asparagine, and leucine). The embryo was characterized by fewer nonadditive patterns, only one low-parent (asparagine) and one high-parent (methionine) patterns have been found. In both cases maternal dominance was the common heterotic pattern.

In endosperm, the levels of most of the amino acids followed a simple, declining time course with few exceptions like aspartate, methionine, glutamate, tyrosine, and lysine (Supplementary Fig. 2). In embryo alanine, glutamate, glycine, leucine, phenylalanine, and tyrosine showed the declining pattern and the majority of the amino acid time courses have a maximum at 20–25 DPP (Supplementary Fig. 3).

Comparison of metabolite patterns

In addition to the description of metabolite time courses and heterotic patterns of individual metabolites, we also investigated whether genotype–time combinations can be distinguished regarding different metabolite patterns. For the purpose of visualizing the different effects of few carbohydrate and many amino acid components on metabolite patterns, hierarchical cluster analysis was carried out separately for these two metabolite classes. In each step those two clusters were joined that had the smallest cluster distance.

The result of the hierarchical cluster analysis of sugar components in whole kernel from 8 to 30 DPP is shown in Fig. 4a. The dendrogram shows distinct groups according to the developmental state of the kernel, representing 8–12, 16–20, and 25–30 DPP. The distances between the three genotypes were smaller than those between developmental stages. The closest individual branches connected either the hybrid or one of the parental lines, preferably B73, in two close developmental states. Similarly, hierarchical cluster analysis of amino acids (Fig. 4b) resulted in two distinct groups according to kernel age. Samples at 8–16 and 20–30 DPP represented the main clusters. Almost all of the closest subclusters were constituted either from B73 and the F1 hybrid or from two Mo17 samples.

Hierarchical cluster analysis of amino acid patterns across the main kernel parts in maize seeds from 16 to 30 DPP resulted in a more complex pattern. Two distinct groups characterized the dendrogram in Fig. 5a. The upper cluster is dominated by endosperm samples, while the lower cluster is dominated by embryo samples, indicating that the majority of the samples are clustered primarily by tissue origin. The few samples that did not cluster by tissue origin in the main cluster were the early endosperm samples in all three genotypes as well as some embryo samples of Mo17 and B73 × Mo17. Within the endosperm and

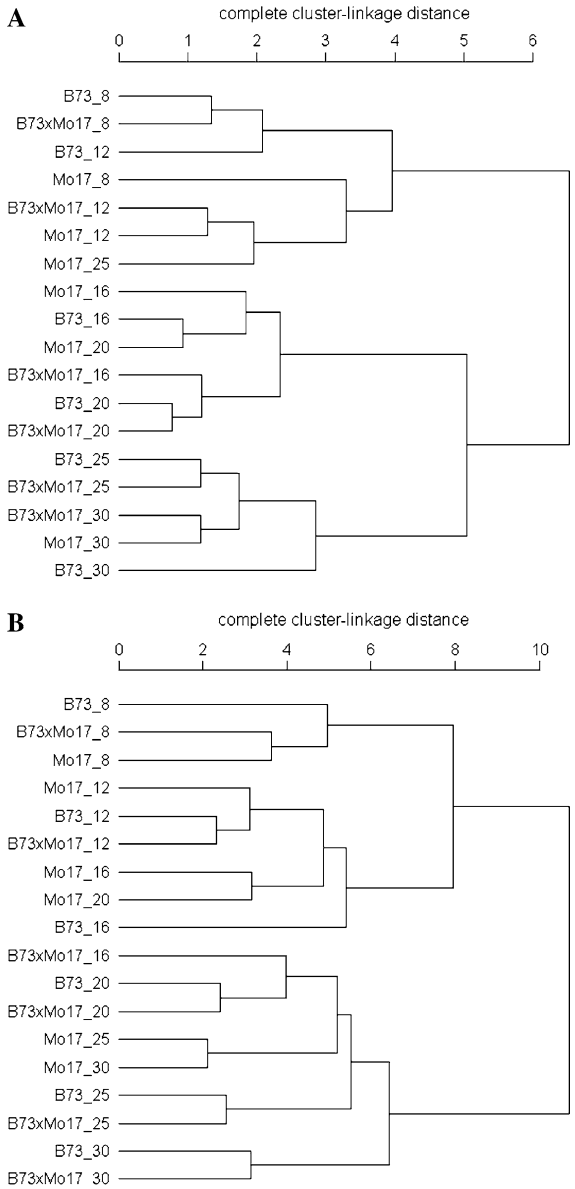

Fig. 4 Hierarchical agglomerative clustering of carbohydrate (**a**) and amino acid (**b**) patterns using complete-linkage agglomeration and Euclidean distances

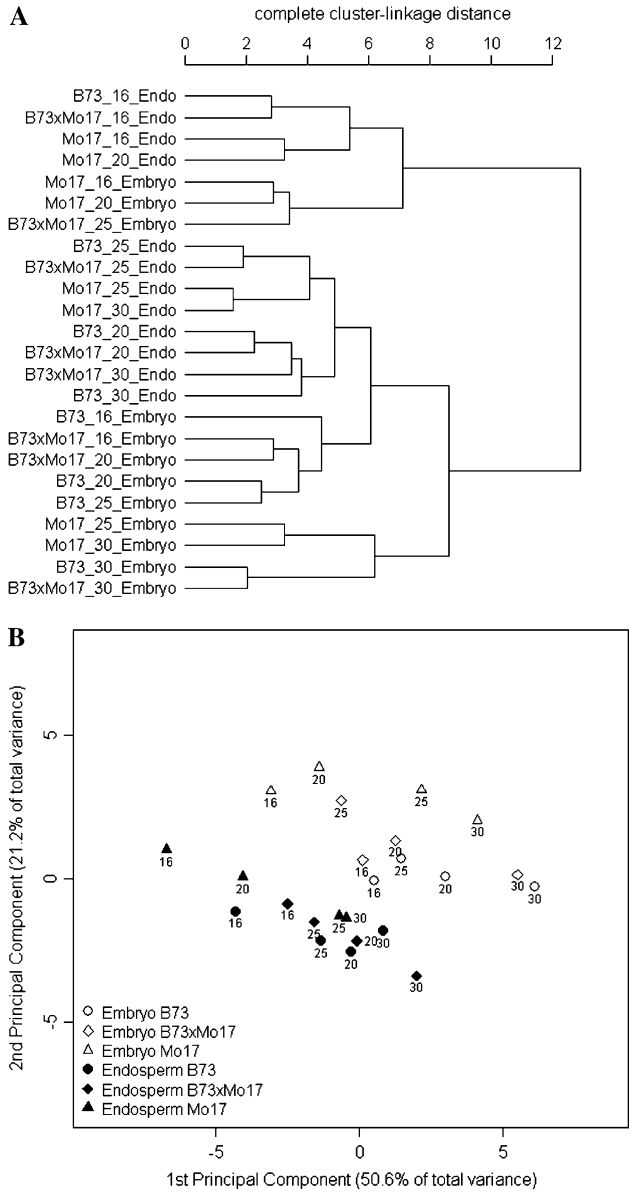

Fig. 5 **a** Hierarchical, agglomerative clustering using amino acid patterns in endosperm and embryo, respectively, at 16–30 DPP. **b** PCA analysis using amino acid patterns in endosperm and embryo, respectively, at 16–30 DPP. The first principal component separates developmental stages; the second principal component separates embryo and endosperm

in the embryo clusters, samples are separated by developmental stages and Mo17 is separated from B73 and B73 × Mo17 lines. The only exception was the B73 × Mo17 embryo sample at 25 DPP. A complementary approach, principal component analysis of the same data set, resulted in a more consistent picture (Fig. 5b). The first principal component resolved amino acid patterns according to developmental stage and the second principal component segregated embryo samples (Fig. 5b, top half) from endosperm samples (Fig. 5b, bottom half). The amino acid composition of Mo17 genotype in both tissue types is distinct from B73 and B73 × Mo17 with regard to the second principal component. Most of the amino acids aligned with the first principal component. Only five amino acids, namely, proline, lysine, histidine, asparagine, and tryptophan align well with the second principal component (data not shown). The amounts of these amino acids showed strong differences between endosperm and embryo tissues as well as between Mo17 and the two other genotypes.

Discussion

Several comparative genome-wide studies in tomato, potato, and *Arabidopsis* (Schauer et al. 2006; Fiehn et al. 2000; Roessner et al. 2000, 2001) demonstrated that metabolite analysis in plants allows the detection of unexpected changes in metabolite levels between homozygous lines and their progeny. We investigated the potential of metabolite profiling to reveal metabolic phenotypes of heterosis in developing maize kernels of the B73 × Mo17 model system. To this end, the sugar and free amino acid compositions as well as starch and protein accumulation were determined in developing kernels between 8 and 30 DPP. We also investigated whether specific patterns of heterosis are observable on the level of individual metabolites.

We found that the abundance of 26 compounds can resolve genotypic differences between B73, Mo17, and B73 × Mo17 kernels. Interestingly, for the majority of the analyzed sugar and amino acid compounds the highest metabolite level was associated with the lowest dry matter and starch content in the Mo17 genotype (Fig. 3). The parental lines, B73 and Mo17, belong to two different heterotic groups, i.e., Iowa stiff stalk synthetic and Lancaster sure crop, respectively. It is therefore not surprising that the frequency of significant differences between B73 and Mo17 was higher than between either parental line and the B73 × Mo17 hybrid. The finding that the metabolic patterns in B73 and hybrid endosperm were more similar to each other than to Mo17 endosperm might results from the two genetic doses of B73 in the triploid endosperm of the hybrid. This hypothesis should be corroborated by the reciprocal cross (Mo17 × B73) that was not part of this study.

For the determination of heterotic effects, the genetic compositions of the investigated tissues were taken into account. Among the metabolites analyzed here, additivity, overdominance, high-parent dominance, low-parent dominance, and underdominance were observed. However, 90% of the metabolite patterns were not different from additivity, e.g., the LS mean determined for the F1 hybrid was not significantly different from the parental average. The majority of the nonadditive heterotic patterns showed low-parent and maternal dominance. Our results are consistent with a global gene expression study (Guo et al. 2003), which indicated that the majority of genes in the triploid endosperm are expressed in an allele-specific additive manner. In this study approximately 8% of the profiled cDNAs were differentially expressed between hybrids of reciprocal crosses and did not fit allelic dosage expression resembling either maternally or paternally expressed genes. There were more genes with maternal-like expression than with paternal-like expression. It could be argued that metabolite concentrations, in contrast to gene expression, are affected in a non-linear way by allelic dosage.

Contrary to the prevalent positive mid-parent heterosis observed in leaves (Lorenz 1972, 1975), in the present study the majority of amino acids as well as glucose and fructose showed negative mid-parent heterosis. This disagreement might be based on the different physiological function and source–sink properties of leaves and kernels. The high abundance of negative mid-parent heterosis as well as the different starch accumulation performance of the genotypes could be related to the study of Meyer et al. (2007a, b), where biomass and the concentrations of most metabolites displayed a negative correlation in *A. thaliana*. It is tempting to speculate that the higher biosynthetic activity of hybrid kernels with respect to the accumulation of reserve compounds causes a more rapid depletion of the carbohydrate precursor pools. Lorenz (1972, 1975) used the differences of seven amino acid concentrations (aspartate, glutamate, glutamine, glycine, alanine, serine, and sarcosine) between parents as biomarkers for hybrid yield in maize. Meyer et al. (2007a, b) reported a negative correlation between biomass and the level of sucrose, glucose, fructose-6-phosphate, glutamine, and phenylalanine. We found that sucrose, asparagine, glutamine, and serine show strong heterotic patterns. These differences with respect to possible marker metabolites among different species and tissues might reflect the different metabolic states of the samples. Therefore, the applicability of metabolites as biomarkers for heterosis has to be validated by inspecting a larger number of inbred–hybrid combinations. However, the first GC–MS based comprehensive study of the heterotic metabolite patterns in *A. thaliana* suggests a complex correlation between heterosis of metabolic and highly integrative phenotypic traits (Lisec et al. 2009).

Our results showed clear changes in the carbohydrate and amino acid patterns during kernel development. These changes likely reflect distinct metabolic programs that satisfy specific demands at major developmental stages of the kernel. We found that heterosis of dry matter biomass and metabolite pattern is already established in the early phase of kernel development, and the relative differences between parental lines and F1 hybrid did not increase towards the middle and late phases of kernel development. This observation is in agreement with an earlier study, which indicated that sink capacity and final kernel weight are established in the early stages of kernel development (Borrás and Westgate 2006).

The amino acid compositions of endosperm and embryo were highly different in each of the three genotypes investigated in this study. This is not surprising given the fundamentally different roles of these tissues in the kernel. There was very little information on free proline levels in

maize kernels prior to the studies of Harrigan et al. (2007a, b), who found high proline concentrations in maize hybrids. Our measurements revealed that the vast majority of the osmoprotectant proline in the kernel is concentrated in the embryo. The prevalence of negative mid-parent heterosis and the frequency of the nonadditive heterotic patterns are similar in both kernel parts, but the distribution of the heterotic patterns is different. Interestingly, a clear interrelationship was not detectable for amino acids between either the time courses, the heterotic patterns, or the tissue specificity and the biosynthetic origin.

In summary, our experiments show that inbred and hybrid genotypes can be distinguished based on the quantification of only a few metabolites. When we compared individual metabolites across genotypes, we predominantly detected higher metabolite concentrations in Mo17. We found that the concentrations of most metabolites in the hybrid were slightly lower than, but not significantly different from, the parental average. Among the nonadditive heterotic patterns, we could detect over- and under-dominance, high- and low-parent dominance as well as maternal and paternal dominance. Most of the metabolites showed low-parent and maternal dominance. The metabolite profiles and the heterotic patterns are specific for endosperm and embryo. The differences in metabolite patterns across the three genotypes occur already early in development. The metabolite composition was a remarkably sensitive indicator for kernel age, which could possibly mask genotypic differences.

Our data show that metabolite profiling can reveal heterotic phenotypes and is complementary to other "omics" technologies, such as profiling of gene expression (Guo et al. 2006), expression of alleles (Guo et al. 2008), and protein expression (Hoecker et al. 2008). Applying metabolite profiling to a series of inbred lines and hybrids in combination with genetic markers is likely to contribute towards understanding and an improved prediction of the heterosis phenomenon in crops.

Acknowledgments We thank Thomas Hoffmann and Florian Vitzthum for support with the GC–MS analysis and Nicole Däschlein for technical assistance. This work was supported by the Deutsche Forschungsgemeinschaft and by the Dr-Ing. Leonhardt-Lorenz-Stiftung.

References

Birchler JA, Yao H, Chudalayandi S (2007) Biological consequences of dosage dependent gene regulatory systems. Biochim Biophys Acta 1769:422–428

Borrás L, Westgate ME (2006) Predicting maize kernel sink capacity early in development. Field Crops Res 95:223–233

Brieger FG (1950) The genetic basis of heterosis in maize. Genetics 35:420–445

Cho K, Torres NL, Subramanyam S, Deepak SA, Sardesai N, Han O, Williams CE, Ishii H, Iwahashi H, Rakwal R (2006) Protein extraction/solubilization protocol for monocot and dicot plant gel-based proteomics. J Plant Biol 49:413–420

Connolly J, Wachendorf M (2001) Developing multisite dynamic models of mixed species plant communities. Annals Bot 88:703–712

Darwin CR (ed) (1876) The effects of cross and self fertilisation in the vegetable kingdom, 1st edn. John Murray, London

Digby PGN, Kempton RA (1987) Multivariate analysis of ecological communities. Chapman and Hall, London

Dobzhansky T (1950) Genetics of natural populations. XIX. Origin of heterosis through natural selection in populations of *Drosophila pseudoobscura*. Genetics 35:288–302

Fiehn O, Kopka J, Dormann P, Altmann T, Trethewey RN, Willmitzer L (2000) Metabolite profiling for plant functional genomics. Nat Biotech 18:1157–1161

Gärtner T, Steinfath M, Andorf S, Lisec J, Meyer RC, Altmann T, Willmitzer L, Selbig J (2009) Improved heterosis prediction by combining information on DNA and metabolic markers. PLoS One 4:e5220. doi:10.1371/journal.pone.0005220.s004

Gower JC, Hand DJ (1996) Biplots. Chapman and Hall, London

Griffing B, Zsiros E (1971) Heterosis associated with genotype-environment interactions. Genetics 68:443–455

Guo M, Rupe MA, Danilevskaya ON, Yang X, Hu Z (2003) Genome-wide mRNA profiling reveals heterochronic allelic variation and a new imprinted gene in hybrid maize endosperm. Plant J 36:30–44

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

Guo M, Yang S, Rupe M, Hu B, Bickel DR, Arthur L, Smith O (2008) Genome-wide allele-specific expression analysis using massively parallel signature sequencing (MPSS) reveals cis- and trans-effects on gene expression in maize hybrid meristem tissue. Plant Mol Biol 66:551–563

Harrigan GG, Stork LAG, Riordan SG, Reynolds TL, Ridley WP, Masucci JD, MacIsaac S, Halls SC, Orth R, Smith RG, Wen L, Brown WE, Welsch M, Riley R, McFarland D, Pandravada A, Glenn KC (2007a) Impact of genetics and environment on nutritional and metabolite components of maize grain. J Agric Food Chem 55:6177–6185

Harrigan GG, Stork LAG, Riordan SG, Ridley WP, MacIsaac S, Halls SC, Orth R, Rau D, Smith RG, Wen L, Brown WE, Riley R, Sun D, Modiano S, Pester T, Lund A, Nelson D (2007b) Metabolite analyses of grain from maize hybrids grown in the United States under drought and watered conditions during the 2002 field season. J Agric Food Chem 55:6169–6176

Hoecker N, Lamkemeyer T, Sarholz B, Paschold A, Fladerer C, Madlung J, Wurster K, Stahl M, Piepho H-P, Nordheim A, Hochholdinger F (2008) Analysis of nonadditive protein accumulation in young primary roots of a maize (*Zea mays* L.) F1-hybrid compared to its parental inbred lines. Proteomics 8:3882–3894

Jahnke S, Sarholz B, Kühr V, Gutiérrez-Marcos J, Geiger H, Piepho H-P, Scholten S, Thiemann A (2010) Heterosis in early seed development: a comparative study of F1 embryo and endosperm tissues six days after fertilization. Theor Appl Genet (in press)

Kenward MG, Roger JH (1997) Small sample inference for fixed effects from restricted maximum likelihood. Biometrics 53:983–997

Kölreuter JG (1766) Vorläufige Nachricht von einigen das Geschlecht der Pflanzen betreffenden Versuchen und Beobachtungen. Engelmann, Leipzig

Lippman ZB, Zamir D (2007) Heterosis: revisiting the magic. Trends Genet 23:60–66

Lisec J, Steinfath M, Meyer RC, Selbig J, Melchinger AE, Willmitzer L, Altmann T (2009) Identification of heterotic metabolite QTL in *Arabidopsis thaliana* RIL and IL populations. Plant J 59:777–788

Lorenz H (1972) Beziehungen zwischen dem Aminosäurepool von Inzuchtlinien und dem Ertrag ihrer Hybriden bei Mais. Z Pflanzenzüchtg 68:155–170

Lorenz H (1975) Relationships between yield and free amino acids in hybrids and inbred lines of corn (*Zea mays* L.). Z Pflanzenzüchtg 75:10–23

Meng F, Ni Z, Wu L, Sun Q (2005) Differential gene expression between cross-fertilized and self-fertilized kernels during the early stages of seed development in maize. Plant Sci 168:23–28

Meyer RC, Törjék O, Becher M, Altmann T (2004) Heterosis of biomass production in *Arabidopsis.* Establishment during early development. Plant Phys 134:1813–1823

Meyer RC, Steinfath M, Lisec J, Becher M, Witucka-Wall H, Toerjek O, Fiehn O, Eckardt A, Willmitzer L, Selbig J, Altmann T (2007a) The metabolic signature related to high plant growth rate in *Arabidopsis thaliana*. Proc Natl Acad Sci USA 104:4759–4764

Meyer S, Pospisil H, Scholten S (2007b) Heterosis associated gene expression in maize embryos 6 days after fertilization exhibits additive, dominant and overdominant pattern. Plant Mol Biol 63:381–391

Moll RH, Lonnquist JH, Fortuno JV, Johnson EC (1965) The relationship of heterosis and genetic divergence in maize. Genetics 52:139–144

Murtagh F (1985) Multidimensional clustering algorithms, in COMPSTAT Lectures 4. Physica Verlag, Wuerzburg

Piepho H-P (2009) Data transformation in statistical analysis of field trials with changing treatment variance. Agron J 101:865–869

Rhodes D, Ju GC, Yang WJ, Samaras Y (1992) Plant metabolism and heterosis. Plant Breed Rev 10:53–91

Roessner U, Wagner C, Kopka J, Trethewey RN, Willmitzer L (2000) Technical advance: simultaneous analysis of metabolites in potato tuber by gas chromatography–mass spectrometry. Plant J 23:131–142

Roessner U, Luedemann A, Brust D, Fiehn O, Linke T, Willmitzer L, Fernie A (2001) Metabolic profiling allows comprehensive phenotyping of genetically or environmentally modified plant systems. Plant Cell 13:11–29

Schauer N, Semel Y, Roessner U, Gur A, Balbo I, Carrari F, Pleban T, Perez-Melis A, Bruedigam C, Kopka J, Willmitzer L, Zamir D, Fernie AR (2006) Comprehensive metabolic profiling and phenotyping of interspecific introgression lines for tomato improvement. Nature Biotech 24:447–454

Shull GF (1908) The composition of a field of maize. Rep Am Breed Assoc 4:296–301

Shull GF (1911) The genotypes of maize. Am Nat 45:234–252

Shull GF (1914) Duplicate genes for capsule form in *Bursa bursa-pastoris*. ZIAV 12:97–149

Shull GF (1952) Beginnings of the heterosis concept. In: Gowen JW (ed) Heterosis. Iowa State College Press, Ames, pp 14–48

Singletary GW, Banisdar R, Keeling PL (1997) Influence of gene dosage on carbohydrate synthesis and enzymic activities in endosperm of starch-deficient mutants of maize. Plant Physiol 113:293–304

Spielbauer G, Margl L, Hannah CL, Gierl A, Bacher A, Eisenreich W, Genschel U (2006) Robustness of central carbohydrate metabolism in developing maize kernels. Phytochemistry 67:1460–1475

Springer NM, Stupar RM (2007) Allelic variation and heterosis in maize: how do two halves make more than a whole? Genome Res 17:264–275

Steinfath M, Gärtner T, Lisec J, Meyer R, Altmann T, Willmitzer L, Selbig J (2010) Prediction of hybrid biomass in *Arabidopsis thaliana* by selected SNP and metabolic markers. Theor Appl Genet (in press)

Stuber CW (1994) Heterosis in plant breeding. Plant Breed Rev 12:227–251

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

Taylor JMG (2006) Transformations–II. In: Kotz S (ed) Encyclopedia of statistical science, vol 14. Wiley, New York, pp 8703–8710

Tsaftaris SA (1995) Molecular aspects of heterosis in plants. Physiol Plant 94:362–370

Wang X, Cao H, Zhang D, Li B, He Y, Li J, Wang S (2007) Relationship between differential gene expression and heterosis during ear development in maize (*Zea mays* L.). J Genet Genomics 34:160–170

Zanoni U, Dudley JW (1989) Comparison of different methods of identifying inbreds useful for improving elite maize hybrids. Crop Sci 29:577–582