Journal of
Experimental
Botany
www.jxb.oxfordjournals.org

RESEARCH PAPER

Variation of metabolic profiles in developing maize kernels up- and down-regulated for the *hda101* gene

Cecilia Castro[1], Mario Motto[2], Vincenzo Rossi[2] and Cesare Manetti[1],*

[1] *Dipartimento di Chimica, Università degli Studi di Roma 'La Sapienza', Piazzale Aldo Moro 5, 00185-Roma, Italia*

[2] *CRA-Unità di Ricerca per la Maiscoltura, Bergamo, Italia*

Abstract

To shed light on the specific contribution of HDA101 in modulating metabolic pathways in the maize seed, changes in the metabolic profiles of kernels obtained from *hda101* mutant plants have been investigated by a metabonomic approach. Dynamic properties of chromatin folding can be mediated by enzymes that modify DNA and histones. The enzymes responsible for the steady-state of histone acetylation are histone acetyltransferase and histone deacetylase (HDA). Therefore, it is interesting to evaluate the effects of up- and down-regulation of a Rpd-3 type HDA on the development of maize seeds in terms of metabolic changes. This has been reached by analysing nuclear magnetic resonance spectra by different chemometrician approaches, such as Orthogonal Projection to Latent Structure-Discriminant Analysis, Parallel Factors Analysis, and Multi-way Partial Least Squares-Discriminant Analysis (N-PLS-DA). In particular, the latter approaches were chosen because they explicitly take time into account, organizing data into a set of slices that refer to different steps of the developing process. The results show the good discriminating capabilities of the N-PLS-DA approach, even if the number of samples ought be increased to obtain better predictive capabilities. However, using this approach, it was possible to show differences in the accumulation of metabolites during development and to highlight the changes occuring in the modified seeds. In particular, the results confirm the role of this gene in cell cycle control.

Key words: Metabolomics, metabonomics, multi-way analysis, NMR, *Rpd3*, *Zea mays*.

Introduction

Eukaryotic genes are regulated by a complex interplay of transcriptional factors and chromatin proteins that pack chromosome DNA into the confined space of the nucleus, while preparing genes for activation or repression (Kadonaga, 1998). Evidence suggests that various levels of chromatin folding ensure the organization of DNA into a tightly packaged environment, which must be highly flexible to switch between repressive condensed to active accessible states. This dynamic property of chromatin can be mediated by enzymes that modify DNA and histones (Fischle *et al*., 2003).

A variety of post-translational modifications of histones has been identified, including acetylation, methylation, phosphorylation, and ubiquitination (Peterson and Laniel, 2004). Among different histone modifications, acetylation of N-terminal lysine residues correlates with transcriptional activation (Shahbazian and Grunstein, 2007). In addition, acetylation is frequently associated with other histone marks, thus forming specific histone modification patterns. These patterns constitute the 'histone code', which is written in response to intra- and extracellular signals, by enzymes that modify histones in specific amino acid residues and is interpreted by regulatory factors that regulate the chromatin structure to modulate gene and genome activity (Jenuwein and Allis, 2001; Berger, 2007).

The enzymes responsible for the steady-state of histone acetylation are histone acetyltransferases (HATs) and histone deacetylases (HDACs). These enzymes are members of distinct gene families and exist as multiprotein complexes (Carrozza *et al*., 2003; Thiagalingam *et al*., 2003). They can be targeted to specific promoters through interaction with sequence-specific transcription factors to modify both histones and non-histone proteins locally and

* To whom correspondence should be addressed: E-mail: c.manetti@caspur.it

can also act globally to modulate the turnover of histone acetylation throughout the genome (Pfluger and Wagner, 2007). HATs, HDACs, as well as other factors involved in the modulation of chromatin structure, are highly conserved in eukaryotes, including plants (Loidl, 2004). However, the peculiarities of plant development and the response to environmental cues can result in marked differences, including the presence of plant-specific HDACs and distinct regulatory mechanisms involved in the establishment and maintenance of epigenetic information. In the genome of different plant species, several potentially functional HDACs have been identified and classified into three distinct families: (i) the Rpd3/Hda1 super-family, (ii) the Sir2-related, and (iii) the plant-specific HD2-like HDACs (http://www.chromdb.org; Pandey *et al.*, 2002). The characterization of Arabidopsis (*Arabidopsis thaliana*) HDAC mutants revealed that members of different classes and within the same group have evolved specific functions (Probst *et al.*, 2004; Tian and Chen, 2001; Zhou *et al.*, 2005; Long *et al.*, 2006). Nevertheless, many aspects of the HDACs' involvement in plant development, as well as the mechanisms responsible for HDAC-mediated control of gene/genome activity, remain elusive.

To date, 15 genes encoding putative HDACs (i.e. 10 *Rpd3/Hda1*-, 1 *Sir2*-, and 4 *HD2-like* genes; http://www.chromdb.org) have been identified in the maize genome, and members of all three HDAC families have been biochemically characterized (Lusser *et al.*, 2001). Studies of maize Rpd3-type HDACs function has revealed that members of this family are differentially expressed during plant development and can physically interact with the maize retinoblastoma-related protein, a key regulator of cell cycle progression (Rossi *et al.*, 2003; Varotto *et al.*, 2003). Furthermore, in cereals, the role of these enzymes in controlling cell division was confirmed by the finding that over-expression of a rice (*Oryza sativa*) *Rpd3* gene leads to alterations in growth rate and plant architecture (Jang *et al.*, 2003). More recently, Rossi *et al.* (2007) used maize plants with specific up- and down-regulation of *hda101* expression to characterize functionally a member of the maize Rpd3-type HDAC family, i.e the *hda101* gene. Their results indicated that gene expression, including the transcription of important regulators of meristem function and of vegetative to reproductive transition, was affected, suggesting a role of *hda101* in modulating plant development, genome activity, and the modification of histone marks. Collectively, the results from the functional characterization of HDA101 indicate that this enzyme affects, either directly or indirectly, the expression of genes involved in various metabolic pathways.

To shed light on the specific contribution of HDA101 in modulating metabolic pathways in the maize seed, changes in the metabolic profiles of kernels obtained from *hda101* mutant plants have been investigated. In the field of metabolomics, the analysis of metabolic changes in time is a fundamental aspect of understanding the biochemical response of an organism to an external perturbation (Lindon *et al.*, 2001). As processes develop through time, the metabolic responses also exhibit dynamic variation. Therefore, monitoring these changes results in characteristic patterns for each type of perturbation. Principal component trajectories have been constructed from Nuclear Magnetic Resonance (NMR) data to investigate the changing multivariate biochemical profile during the development of a toxic lesion (Keun *et al.*, 2004). However, this kind of analysis, although effective for trajectory analysis, is not suitable for the simultaneous comparison of several parallel systems, and thus the use of alternate multi-way tools for optimally extracting metabolic trajectory and biomarker information have been investigated (Antti *et al.*, 2002; Dyrby *et al.*, 2005a). Multi-way analysis (Bro, 1997) of NMR data is the extension of the traditional multivariate analysis, already applied to metabonomic studies of maize (Manetti *et al.*, 2004, 2006) and permits the direct study of development through time.

Among the existing procedures, several approaches were used in this study to distill information from the large amount of experimental data. As a first step, PARallel FACtor analysis (PARAFAC) was chosen (Bro, 1997); this is a multi-way decomposition method, which can be compared to bi-linear Principal Component Analysis (PCA), or rather it is *one* generalization of bi-linear PCA.

The important difference between PCA and PARAFAC is that in PARAFAC there is no need for orthogonality to identify the model. Furthermore, the PARAFAC model has the advantage of being a unique solution (Bro, 1997), an important characteristic when the dynamics of changes are studied. PARAFAC was applied successfully in many areas, ranging from the monitoring of the photocatalytic degradation of phenol in aqueous suspensions of TiO_2 by fluorescence spectroscopy (Bosco *et al.*, 2006) to the quantification of nitrite in water and meat samples by kinetic-spectrophotometric analysis (Niazi *et al.*, 2005), and to the prediction of sensory qualities of different potatoes (Povlsen *et al.*, 2003).

Subsequently, multi-way Partial Least-Squares-Discriminant Analysis (N-PLS-DA) was applied; this is the regression method for the analysis of a higher order array (Bro, 1996). As for the traditional two-way PLS, it searches for a compromise between better fit (i.e. less error in describing the array of the independent variables) and better prediction (i.e. less error in evaluating the response space). N-PLS has been applied successfully in many areas, ranging from the analysis of food characteristics by fluorescence (Christensen *et al.*, 2006) and gas chromatography (Guimet *et al.*, 2005; Durante *et al.*, 2006), to the simultaneous determination of ions by

electrochemical sensors (Chow *et al.*, 2006). Recently, the use of N-PLS in the quantification of lipo-protein fractions obtained by 2-D diffusion-edited NMR spectra has been evaluated (Dyrby *et al.*, 2005*b*).

In summary, in this paper, the application of multi-way methods to NMR spectra to evaluate the effect of the up- and down-regulation of HDA101 activity in terms of metabolite concentrations during maize seed development is described, comparing the two approaches one with the other. In addition, the results are compared with those of Orthogonal Projection to Latent Structure-Discriminant Analysis (OPLS-DA), an abundantly used bivariate approach in the metabonomic field (Cloarec *et al.*, 2005*a*; Trygg and Wold, 2002).

Materials and methods

Plant material

Kernel samples from the B73 inbred lines (WT) and two early isogenic versions of this inbred, containing a modified *ZmRpd3-101* maize gene (Rossi *et al.*, 1998), in sense and antisense orientation (OE1 and AS33), respectively, were used in this study. A detailed description of the origin of the transgenic lines was recently reported by Rossi *et al.* (2007) and a summary will be given here. Briefly, the constitutive maize ubiquitin promoter (Christensen *et al.*, 1992) was cloned upstream or downstream of the full-length *hda101* cDNA sequence, into the pGEM3-ZmRpd31 plasmid (Rossi *et al.*, 2003), to obtain sense and antisense *hda101* constructs, respectively. The polyadenylation domain of the *nopa- line synthase* gene was inserted opposite to the *ubiquitin* promoter. The resulting cassettes were used to generate the pRpd3-53 and pRpd3-35 plasmids. These plasmids were employed to transform protoplasts from the maize suspension cell line HE-89 (Morocz *et al.*, 1990) using the PEG method. Regenerated T_0 plants were converted to the B73 inbred by two backcrosses, to minimize a mixed genetic background influence, and selfed twice to obtain homozygous plants, which were used for analysis. The presence of the transgenes in plants was assessed for resistance to gluphosinate using PCR screening.

For the current study two lines were chosen, i.e. AS33 and OE1, which displayed the most pronounced differences in the levels of both *hda101* mRNA and protein compared with the wild type; differences in *hda101* expression was observed in seedlings, developing ears, and kernels harvested at different developmental stages (Rossi *et al.*, 2007).

Plants of the inbred line B73 and its transgenic versions were grown in experimental plots under containment, according to guidelines of the Italian laws for bio-safety. At flowering, plants were self-pollinated; and a minimum of six well-filled ears of each genotype were harvested at four stages of kernel development, i.e. 8, 13, 18, and 23 days after pollination (DAP) and frozen immediately in liquid N_2. Kernels harvested at physiological maturity were also used in this study. For each genotype, the kernel samples were conserved in sealed plastic bags at –80 °C until NMR analysis.

NMR sample preparation

For each genotype, a sample of single maize seed was weighed and then frozen in a stainless steel mortar using liquid N_2, before being pulverized to a fine powder with a pestle chilled in liquid N_2 and maintained in a liquid N_2 bath during the pulverization procedure.

Three ml of methanol/chloroform mixture (2:1 v/v) were added to the entire pulverizated sample. The powder was stirred and 1 ml of chloroform and 1.2 ml of water were added (Bligh-Dyer modified; Miccheli *et al.*, 1988; Ricciolini *et al.*, 1994). The sample was stored at 4 °C for 1 h and then centrifuged at 10 000 *g* for 20 min at 4 °C. The resulting upper hydro-alcoholic and lower chloroformic phases were separated. The extraction procedure was performed twice, once on the powder and once on the pellet, in order to obtain a quantitative extraction. After the second extraction, the two hydro-alcoholic phases obtained were pooled, dried under N_2 flux, and stored at –80 °C prior to analysis.

NMR data collection

To obtain NMR spectra, the dried sample was dissolved in 1 ml of 0.5 mM TSP (sodium salt of 3-(trimethylsilyl)propionic-2,2,3,3-d_4 acid) solution in D_2O PBS buffer (pH 7.4) to avoid chemical-shift changes due to pH variation. The dissolved extracts were trans- ferred to a 5 mm NMR tube.

NMR spectra were recorded on a Bruker (Bruker GmbH, Rheinstetten, Germany) DRX 500 spectrometer, operating at ^{1}H frequency of 500.13 MHz. Single pulse spectra were acquired using a solvent pre-saturation pulse sequence to suppress residual water resonances, so that it was possible to evaluate signals near the solvent region better. Possible bias was taken into account in the post-processing procedure, eliminating variables (i.e. buckets, see below) corresponding to the water region. Spectra were obtained at $T=27$ °C, 256 scans were acquired, with data collected into 64 k data points, and a spectral width of 12 ppm, using a 20 s delay for a full relaxation condition. Prior to Fourier transformation, an exponential multiplication was performed, using a line broaden- ing equal to 0.09 Hz.

The spectra were phased, baseline corrected using the usual ACD routines (Advanced Chemistry Development Inc., 90 Adelaide Street West, Toronto, Ontario, M5H 3V9, Canada), and they were referenced to TSP for chemical shift (0.00 ppm).

NMR data pre-processing treatment

The ^{1}H spectra were reduced to 499 buckets to produce a matrix of sequentially integrated regions of 0.02 ppm in width between –0.5 ppm and 9.5 ppm, using ACD software: column 1 corresponds to the bucket –0.5 ppm to –0.48 ppm. The spectra were normalized according to the area of the TSP peak, fixed to 10. The region between 4.7 ppm and 4.9 ppm was removed to eliminate baseline effects due to presaturation of the water signal.

Statistical analysis

Orthogonal Projection to Latent Structure-Discriminant Analysis (OPLS-DA): The reduced and normalized NMR spectral data were inported into Matlab (version 7.4, The Mathworks, Natick, MA).

OPLS-DA has gained increasing interest in the metabonomic field, due to the possibility of obtaining models that allow a thorough interpretation of the results. This is achieved by a separate modelling of predictive (variation of interest) and class- related (variation not related to the responses) variation in the X-matrix (the spectral descriptors being metabolite concentrations or buckets) through the identification of Y-orthogonal variation (Bylesjö *et al.*, 2007) (Y being the matrix made of the features of interest such as treatment classes).

In this case, X has been considered as a matrix containing the bucketed spectra and Y as a dummy matrix containing information on the seed genotype. In particular, this input matrix has a row, for

each sample, containing 1 for the y variable corresponding to the right group and 0 for all the others.

Data have been mean centred and unit variance scaled before investigation and analysed using in-house routines. Note that this approach collapses the time dimension by concatenating the NMR spectra. In this way, explicit reference to time is lost while this dimension can be crucial to study a developmental process.

Multi-way analysis: In order to include the time dimension, data were arranged in a three-way data box with kernels as mode 1, bucketed NMR spectra as mode 2, and time as mode 3, giving a data cube of dimension 12×484×5, as it consists of 4 replicates×3 genotypes×5 growth stages as rows and 484 NMR regions as variables.

Matlab was equipped with the N-way Toolbox version 3.0 (obtained from R Bro at http://www.models.life.ku.dk/source; Andersson and Bro, 2000).

Prior to analysis, the data were mean centred and scaled at unit variance across the sample mode.

Initially, data were analysed by an unsupervised multi-way approach, such as PARallel FACtor analysis (PARAFAC), where decomposition of the data is made into triads or trilinear components. Instead of one score vector and one loading vector as in bilinear PCA, each component consists of one score vector and two loading vectors. A PARAFAC model of a three-way array is given by three loading matrices, A, B, and C with typical elements a_{if}, b_{jf}, and c_{kf} and it is defined by the structural model

$$\hat{x}_{ijk} = \sum_{f=1}^{F} a_{if} b_{jf} c_{kf}$$

The second multi-way approach used was multi-way partial least-squares (N-PLS), which allows the regression of NMR data against classes present in the data set. For this procedure also, time was explicitly included and data were arranged in a box, putting behind each other data sheets ordered in time. A Y matrix was constructed, where each column defines a group that corresponds to the genotypes of the kernels, whose values are dummy variables. In particular, the column contains '1' for the samples belonging to the group and zeros for all the others. The method could therefore be defined as N-PLS-DA. Figure 1 shows a schematic representation of the strategy that has been applied here.

As a result of the N-PLS-DA model, the X array is decomposed in terms of a scores matrix (T) relative to the sample mode, and two weights matrices (W_J, W_K), relative to the NMR and time modes, respectively, while the Y array is decomposed in terms of a score matrix (U) relative to the first mode, and one loadings matrix Q_M relative to the second mode. The expression relating the two decomposition models of X and Y is:
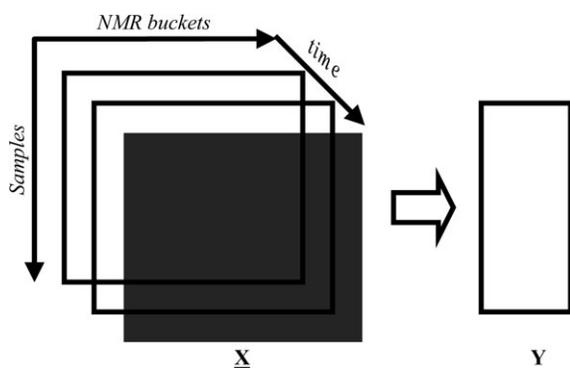
$$U = TB + E_U,$$

where B represents the regression coefficients, while E_U the error matrix.

Prior to analysis, the data were mean centred and scaled at unit variance across the sample mode. Following the approach applied by Durante *et al.* (2006), the model was validated using a leave-one-genotype-out-at-a-time: samples belonging to the same genotypes were always left in the same validation segment.

Results

Transgenic plants with up- and down-regulation of hda101 expression

Maize plants with up- and down-regulation of *hda101* transcription were generated by transformation with plasmids that constitutively over-express *hda101* sense (OE) or antisense (AS) transcript (Rossi *et al.*, 2007). A phenotypic characterization of the two transgenic progenies of plants revealed that the perturbation of *hda101* expression induces alterations in plant growth and development, including kernel size (Table 1). In particular, the kernel size, measured as the 100-kernel weight, was definitively smaller in OE1 mature kernels (16.5±0.4 g) compared with wild-type (27.9±1.4 g) and AS33 (22.5±1.0 g).

Evaluation of ^{1}H NMR spectra of maize seed

In the current study, the ^{1}H-NMR spectra of hydro-alcoholic extracts of maize kernels derived from AS33, OE1, and WT lines and harvested at five different developmental stages have been analysed. The stages investigated reflect: (i) 8 DAP, corresponding to the embryogenesis stage, when the embryo differentiates; (ii) 13 DAP, corresponding to the initial stage of maturation characterized by cell expansion; (iii) 18 DAP, corresponding to the central phase of kernel maturation when the maximum accumulation of starch and storage proteins occurred; (iv) 23 DAP, corresponding to the end phase of maturation, when their water content starts to decrease; and (v) mature seeds. The assignment of the peaks was achieved in data based on the literature (Fan, 1996; Le Gall *et al.*, 2003) and on 2D spectra, previously obtained for maize seed samples in our laboratory (Manetti *et al.*, 2006).

In Fig. 2, NMR spectra for all the developmental stages of the WT seeds are reported. The list of the assigned metabolites is reported in Table 2. A number of assignable

Fig. 1. Schematization of the application of the N-PLS-DA method.

Table 1. *Means and standard errors for morphological plant traits of B73 parental line (wild type), and AS33 and OE1 lines*

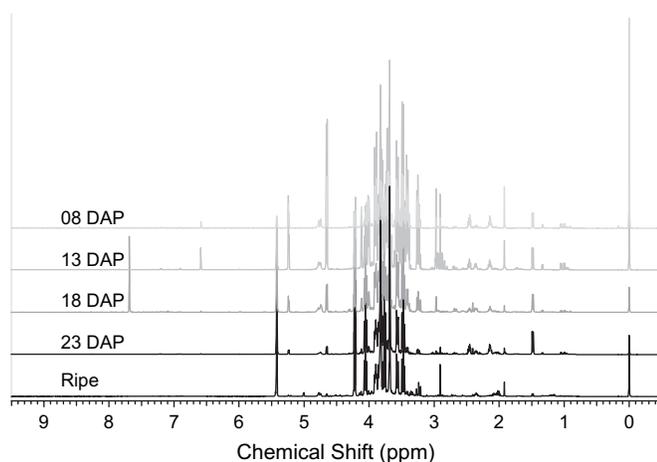| Traits	wt	AS-33	OE1
Seedling dw (mg)	187±33	140±23	60±4
Pollen shed (GDD)	842±13	931±18	998±32
Plant height (cm)	186±12	150±16	135±8
100 kernel weight (g)	28±1	22.5±1.0	16.5±0.4

Fig. 2. Plot of NMR spectra for all the developmental stages of the control seeds.

amino acids, organic acids, and sugars were identified. In particular, the following metabolites dominated the spectra: (i) amino acids such as threonine, alanine, glutamate, glutamine, aspartate, asparagine, tyrosine, phenylalanine, valine, isoleucine, and acid γ-amino butyric acid (GABA); (ii) organic acids such as pyruvate, succinate, 3-hydroxybutyrate, and choline; and (iii) sugars such as α- and β-glucose and sucrose. Significant ratios compared with the wild type were observed in the transgenic kernels for several metabolites during the different stages of sampling. The differences of the metabolite levels in transgenic compared to wild-type kernels are shown in Table 3 and Table 4 and are particularly evident in mature kernels. Specifically, down-regulation of *hda101 (*AS33 line) gave rise to a reduction in the level of several metabolites (Table 3). In fact, these seeds were associated with a decrease in the level of sugars, organic acids such as acetate, 3-hydroxybutyric acid, and succinate, and amino acids such as asparagine, glutamine, isoleucine, valine, and GABA. Considering the TCA-cycle intermediates, succinate was the only metabolite that displayed a major reduction in content. Kernels derived from plants with up-regulation of *hda1* (OE1 line) were generally associated with a significant accumulation of the metabolite concentrations with respect to wild type, ranging from 1.3–4.3-fold, with the exception of 3-hydroxybutyric acid, acetate, and formate for which the ratios, are <1 (Table 4). In detail, there was a significant accumulation in the content of most of the amino acids and, more notably, in the non-proteogenic amino acid GABA. Considering glycolysis and the TCA-cycle, the intermediates pyruvate, succinate, and malate, were the only metabolites that displayed a considerable increase in their content. The amount of many metabolites was also altered in AS33 and OE1 kernels harvested at different developmental stages, although the number of assignable metabolites showing variation is smaller than the one observed in mature kernels.

Analysis by OPLS-DA

To extend the knowledge obtained by univariate analysis (i.e. variations in levels of single metabolites) on the dynamics of processes and on differences among genotypes, a metabonomic approach was applied. As a first step, data were investigated by OPLS-DA, to remove variations in X (NMR spectra) unrelated to Y (genotype). The resulting model was constituted by ten latent variables explaining 66.12% of X-variance and 72.72% of Y-variance. All three genotypes formed distinct clusters in the space spanned by the scores of the first two latent variables (Fig. 3). Furthermore, the orthogonal matrix obtained by the analysis was investigated by PCA: it contains characteristics correlated to the developing process, as shown by the score plot (Fig. 4). The samples are ordered according to their development stage on the second principal component, even if the data matrix does not explicitly take time information into account.

Analysis by multi-way methods

To extend the analysis, a multi-way approach was chosen to investigate the data set by considering the characteristics of the system (Castro and Manetti, 2007), containing two factors that simultaneously change (i.e. genotype and kernel development).

PARAFAC analysis was carried out on the array containing the NMR bucketed spectra. As a result, a model with two latent variables was obtained, explaining 67.25% of variance. In Fig. 5, a score plot is shown. It is noticeable that both factors manage to discriminate between OE1 derived kernels and those of the other two genotypes.

N-PLS-DA regression was developed between the array containing the bucketed NMR spectra and a 'dummy matrix' containing information about samples groups. As a result, two factors were found to be significant, explaining 46.33% of X variance and 77.63% of Y variance. The total explained variation of X for this model is relatively low because many regions of the spectra contained only noise, and the autoscaling of the corresponding variables contributes to increasing the random variance, which is impossible to model. However, 76% of Y variance can be related to this variation in X.

In Fig. 6, the score plot is shown. It is evident that it is possible to separate the three genotypes. In particular, it is noticeable that the first factor manages to discriminate between OE1 derived kernels, characterized by positive values of the scores, and those of the other two genotypes, with negative values. Furthermore, the second factor allows the discrimination between the wild-type kernels, which have positive values, and the two genetically modified kernels with negative values.

In Fig. 7, the plot of the loadings corresponding to the NMR mode is shown. For the first latent variable, positive

Table 2. *Resonance assignments of metabolites identified in NMR spectra of maize seeds at the different stages of development*

δ (ppm)	Multiplicity[a]	Assignment	Stage of development
0.94	t	Ile (δ-CH$_3$)	All
0.95	d	Leu (δ′-CH$_3$)	All
0.97	d	Leu (δ-CH$_3$)	All
0.99	d	Val (γ-CH$_3$)	All
1.01	d	Ile (γ′-CH$_3$)	All
1.05	d	Val (γ′-CH$_3$)	All
1.14	d	Isobutyrate (CH$_3$)	18, 23 DAP, Ripe
1.21	d	3-Hydroxybutyrate (γ-CH$_3$)	Ripe
1.33	d	Thr (γ-CH$_3$)	All
1.48	d	Ala (β-CH$_3$)	All
1.72	m	Leu (β-CH$_2$)	All
1.9	q	GABA (β-CH$_2$)	All
1.92	s	Acetate	All
2.02	m	Pro (γ-CH$_2$)	23 DAP, Ripe
2.04	m	Pro (β′-CH)	23 DAP, Ripe
2.05	m	Glu (β-CH$_2$)	All
2.14	m	Gln (β-CH$_2$)	All
2.3	t	GABA (α-CH$_2$)	All
2.36	m	Glu (γ-CH$_2$)	All
2.36	m	Pro (β-CH)	23 DAP, Ripe
2.4	s	Pyruvate	13 (AS33), 18, 23 DAP, Ripe
2.41	s	Succinate	13 (AS33), 18, 23 DAP, Ripe
2.46	m	Gln (γ-CH$_2$)	All
2.67	dd	Asp (β-CH)	All
2.73	s	Dimethylamine	13 (AS33), Ripe (C4)
2.82	dd	Asp (β′-CH)	All
2.86	dd	Asn (β-CH)	All
2.91	s	Trimethylamine	All
2.92	dd	Asn (β-CH)	All
3.02	t	GABA (γ-CH$_2$)	All
3.15	dd	Tyr (β-CH)	08, 13, 18, 23 DAP
3.21	s	Choline (N-CH$_3$)	All
3.25	dd	β-Glc (C2H)	08, 13, 18, 23 DAP
3.33	t	Pro (δ′-CH)	23 DAP, Ripe
3.4	dd	β-Glc (C4H)	All
3.42	dd	α-Glc (C4H)	All
3.47	t	β-Glc (C5H)	08, 13, 18, 23 DAP
3.49	t	β-Glc (C3H)	08, 13, 18, 23 DAP
3.49	t	Sucrose (G4H)	All
3.52	dd	α-Glc (C2H)	All
3.58	dd	Sucrose (G2H)	All
3.68	s	Sucrose (F1H)	All
3.7	t	α-Glc (C3H)	08, 13, 18, 23 DAP
3.72	dd	α-Glc (half-C6H)	08, 13, 18, 23 DAP
3.77	m	α-Glc (half-C6H)	08, 13, 18 DAP
3.78	t	Sucrose (G3H)	All
3.82	m	β-Glc (half-C6H)	08, 13, 18 DAP
3.83	dd	α-Glc (C5H)	08, 13, 18 DAP
3.83		Sucrose (G6H)	All
3.84		Sucrose (G5H)	All
3.84	c	Sucrose (F6H)	All
3.89	dd	β-Glc (half-C6H)	08, 13, 18 DAP
3.91	m	Sucrose (F5H)	All
4.01		α-Fructose (C4H)	08, 13, 18, 23 DAP
4.04	c	α-Fructose (C5H)	08, 13, 18, 23 DAP
4.06	t	Sucrose (F4H)	All
4.12		Fructose (C3H)	08, 13, 18, 23 DAP
4.22	d	Sucrose (F3H)	All
4.3	dd	Malate (α-CH)	All
4.65	d	β-Glc (C1H)	All
5.24	d	α-Glc (C1H)	All

Table 2. *Continued*

δ (ppm)	Multiplicity[a]	Assignment	Stage of development
5.42	d	Sucrose (G1H)	All
5.58	d	Glc 6-P	13, 18, 23 DAP, Ripe
5.97	d	UMP (C5 ring)	23 DAP
5.99	d	UMP (C1′ H ribose)	23 DAP
6.52	s	Fumarate	18, 23 DAP
6.9	d	Tyr (C3,H5 ring)	All
7.2	d	Tyr (C3, H5 ring)	All
7.4	m	Phe	08 (OE1), 13, 23 DAP, Ripe
7.69	s	Guanine	08, 13, 18 DAP
8.09		Trigonelline (HB, HC)	Ripe
8.11	d	UMP (C6 ring)	23 DAP
8.46	s	Formate	All
8.85		Trigonelline (HB, HC)	Ripe
9.13		Trigonelline (HA)	Ripe

[a] s, Singlet; d, doublet; dd, doublet of doublets; t, triplet; q, quartet; m, multiplet; c, complex.

values are obtained for NMR descriptors, corresponding to all the resonances of sucrose, α and β-glucose, and choline. For the second latent variable, alanine, glutamine, and sucrose have positive loading values, while α and β-glucose and guanine have negative ones.

The efficacy of the method in describing a dynamic trajectory is shown in Fig. 8, where the loadings corresponding to the time mode are plotted; this analysis shows the discriminant trajectories among the genotypes and this is easily correlatable to the evolution of the sysyem in time as seen by NMR spectra (Fig. 2).

Discussion

In this study the metabolic profiles of three maize genotypes were compared to highlight the consequences of the up- and down- regulation of *hda101* gene in developing kernels. The discrimination among the genotypes was obtained, taking into account the growth pattern of the three groups.

Data have been analysed by different multivariate analysis methods. In particular, N-PLS-DA has been applied, considering as the response block a 'dummy matrix', which records samples membership, and obtaining a model based on the discriminating characteristics among the groups analogous to the PLS-Discriminant Analysis (PLS-DA) approach, that is abundantly used in metabonomics. The advantage of a multi-way analysis in comparison to the two-way multivariate technique is not to obtain a better fit, but rather to obtain more adequate, robust, and interpretable models (Bro, 1997). In fact, the constructed models integrate the information contained in the entire structure of the multi-way array, without any loss due to the unfolding procedure.

Table 3. *Results (ratios from AS-33 to wild-type) from selected signals from maize spectra at different stages of kernel development*

Compound	ppm	Stage of kernel development (DAP)[a]				
		8	13	18	23	Mature
3-Hydroxybutyric acid	1.14..1.14	0	3.116	0.625	9.749***	0.433**
GABA	2.28..2.29	0.866	0.926	0.378	0	0.732*
Acetate	1.87..1.94	0.596*	0.784	0.480***	0.681**	0.941
Alanine	1.47..1.50	0.308*	1.240	1.018	0.552	1.097
Asparagine	2.85..2.86	1.129	0.685	1.908	1.263	1.328*
Isoleucine	1.02..1.03	0.482*	0.637*	0.872	0.901	0.481***
Valine	1.03..1.05	0.532*	0.882	0.744*	0.745	0.676**
Tyrosine	6.89..6.92	0.997	0.658**	0.905	0.536	1.107
α-Glucose	5.23..5.23	0.967	0.629***	1.235	1.104	0.558*
β-Glucose	4.63..4.67	0.916	0.613**	1.186*	0.978	0.456**
Choline	3.20..3.22	0.867	1.497*	2.312*	1.363	1.077
Formate	8.45..8.47	0.517	0.710	0.805	1.038	1.060
Histidine	7.83..7.85	2.807*	1.392	1.059	2.029*	0.903
Succinate	2.40..2.41	0.586	2.635*	0.746*	1.517	0.410***
Pyruvate	2.39..2.40	0.915	3.062*	0.988	1.530	0.435**
Malate	4.29..4.32	0.973	1.872	0.990	1.515*	1.017
Sucrose	3.67..3.70	0.906	1.228	0.832**	0.736**	1.040
Threonine	1.32..1.35	0.460**	0.889	0.774*	0.800	1.130
Glutamine	2.44..2.49	0.563	1.246	0.771	0.709	0.735***

[a] *, **, *** Represent statistically significant difference at the 0.05, 0.01, 0.001 probability levels, respectively.

Table 4. *Results (ratios from OE1 to wild type) from selected signals from maize spectra at different stages of kernel development*

Compound	ppm	Stage of kernel development (DAP)[a]				
		8	13	18	23	Mature
3-Hydroxybutyric acid	1.14..1.14	0.169	0.700	0.878	0.397*	0.549*
GABA	2.28..2.29	0.964	1.213	0.679	0.348	4.224***
Acetate	1.87..1.94	0.535*	0.730	0.764**	0.706***	0.737*
Alanine	1.47..1.50	0.253**	0.804	0.771*	0.805	3.065***
Asparagine	2.85..2.86	0.236***	0.664	0.821	0.268	1.342*
Isoleucine	1.02..1.03	0.656	1.251	0.638	1.105	1.494**
Valine	1.03..1.05	0.539*	0.800	0.704**	0.743	1.674*
Tyrosine	6.89..6.92	1.169	0.847	0.672***	0.415*	2.068***
α-Glucose	5.23..5.23	1.087	0.826*	1.158	1.167	2.766***
β-Glucose	4.63..4.67	0.915	0.760*	1.198	1.096	2.535***
Choline	3.20..3.22	1.291	0.958	2.014	2.095	1.830***
Formate	8.45..8.47	0.471	0.673*	0.936	0.541***	0.738*
Histidine	7.83..7.85	0.779	0.438*	0.957	1.663	1.841**
Succinate	2.40..2.41	0.394*	0	0.631**	0.889	2.748***
Pyruvate	2.39..2.40	0.792	0	0.741*	1.347	1.666*
Malate	4.29..4.32	0.921	0.612	0.619*	0.969	1.503*
Sucrose	3.67..3.70	0.965	0.271	0.730**	0.657	1.314*
Threonine	1.32..1.35	0.583**	0.909	0.832*	1.296	2.033*
Glutamine	2.44..2.49	0.234**	0.415	0.387*	0.297	1.204*

[a] *, **, *** Represent statistically significant difference at the 0.05, 0.01, 0.001 probability levels, respectively.

Our aim is the analysis of changes in the developing process due to a genetic modification perturbing the expression of genes involved in various metabolic pathways by a change in the organization of histone.

The chosen approach is based on the study of metabolism through different stages of development by an NMR technique, whose non-targeted characteristics allow the observation of variations on the entire bucketed spectra with hundreds variables for each spectra. The comparison between variable and sample numbers could suggest that it is impossible to build a model suitable for prediction and not only for summarizing huge data sets. However, it is crucial to keep in mind that high correlation exists among many variables (spectrum buckets). In fact, NMR sensibility limits the observable metabolites to some tens, and the spectra are constituted by many signals arising from different parts of the same molecule, signals that are a duplication of information. In fact, correlation analysis of these signals, highlighted by specific pulse sequences (COrrelation SpectroscopY, COSY; Total
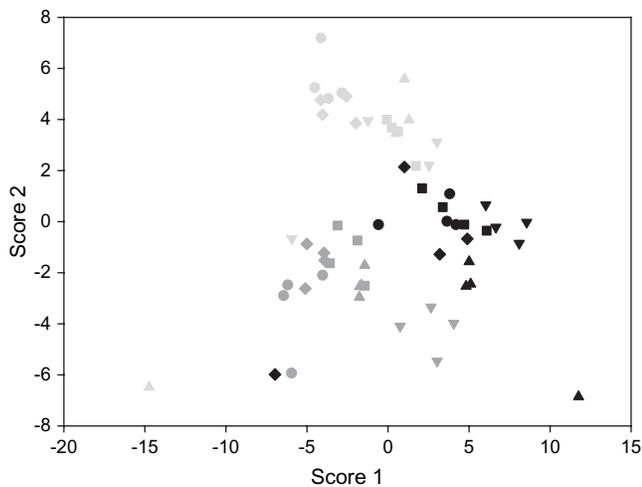
Fig. 3. Score plot obtained by OPLS-DA. Seeds collected at 8 DAP are represented as circles, those collected at 13 DAP as squares, those collected at 18 DAP as diamonds, those collected at 23 DAP as triangles and the ripe seeds as inverted triangles. Black symbols refer to wild-type seeds, grey ones refer to the antisense-mediated down-regulation of the *hda101* gene, and dark grey symbols refer to the over-expression of the *hda101* gene.
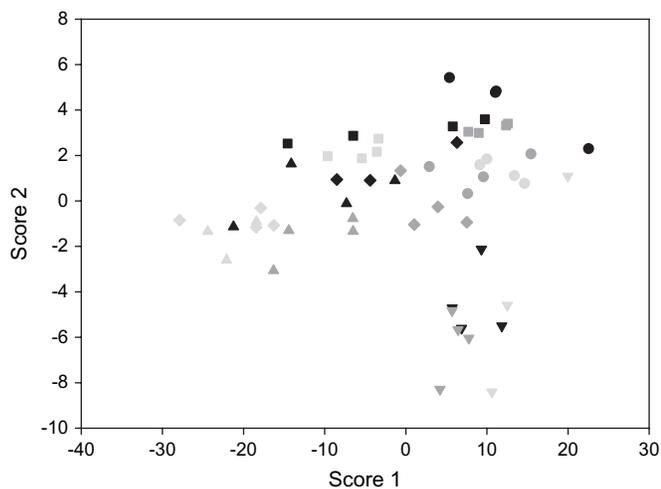

Fig. 4. Score plot obtained by PCA performed on the ortogonal matrix resulting from OPLS-DA. Symbols and colours are the same used in Fig. 2.

Fig. 5. Score plot obtained by N-PLS-DA. The wild-type samples are represented as C, antisense-mediated down-regulation of the *hda101* gene (AS33) as AS, and over-expression of the *hda101* gene (OE1) as OE. The ellipse represents the Hotelling T^2 with 95% confidence.

Fig. 6. Score plot obtained by N-PLS-DA. The wild-type samples are represented as C, antisense-mediated down-regulation of *hda101* gene (AS33) as AS, and over-expression of the *hda101* gene (OE1) as OE. The ellipse represents the Hotelling T^2 with 95% confidence.

Correlation SpectroscopY, TOCSY), allows spectral assignment: this purely spectroscopic approach found a chemometrician equivalent (Statistical TOtal Correlation SpectroscpY, STOCSY) when the metabolomic approach tried to highlight correlations among metabolites due to system biochemistry, i.e. metabolic pathways (Cloarec *et al.*, 2005*b*).

At the beginning, the metabolomic approach was mainly focused on classification problems, i.e. the characterization of healthy versus ill patients (score plot analysis) (Nicholson *et al.*, 2002), nowadays, this approach is extended to metabolic trajectories study by the
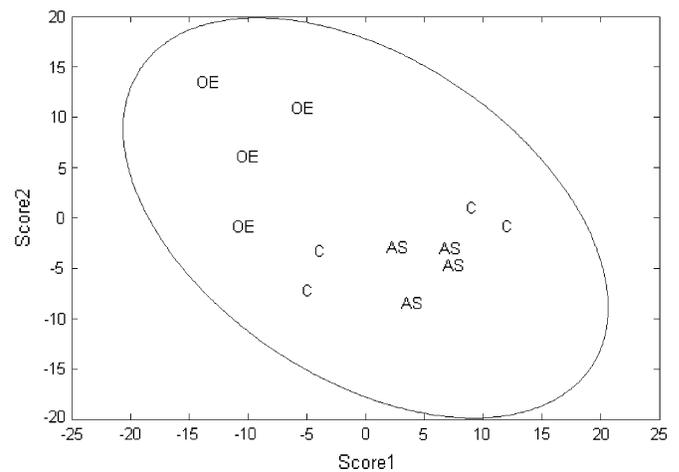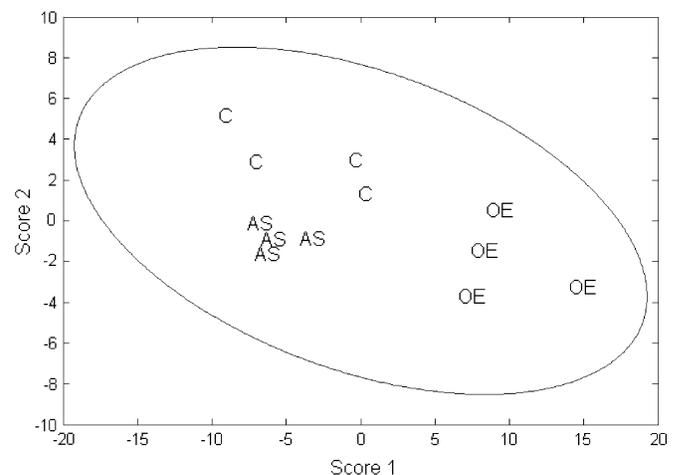
analysis of original variable contributions to the definition of the new latent variables (loading analysis) driving the observer to go in-depth to the metabolic significance of data (Manetti *et al.*, 2006; Coen *et al.*, 2008).

This second aim is less correlated to the predictive capabilities of the model, but it highlights some of the discriminant characteristics among sample classes. The classification obtained can be driven by multiple regression techniques (e.g. PLS-DA) that give the possibility of observing the correlation structure from the desired point of view. Analysis of the loading obtained (Fig. 7) allows correlation among metabolites: this correlation is not always direct, but can be guaranteed by the metabolites involved in other pathways. For these reasons, in
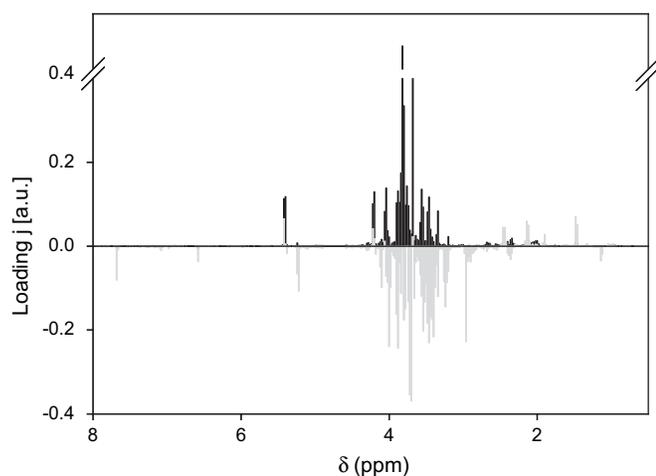
Fig. 7. Plot of the loadings along the ppm dimension. The black curve refers to the first latent variable, while the grey one to the second.
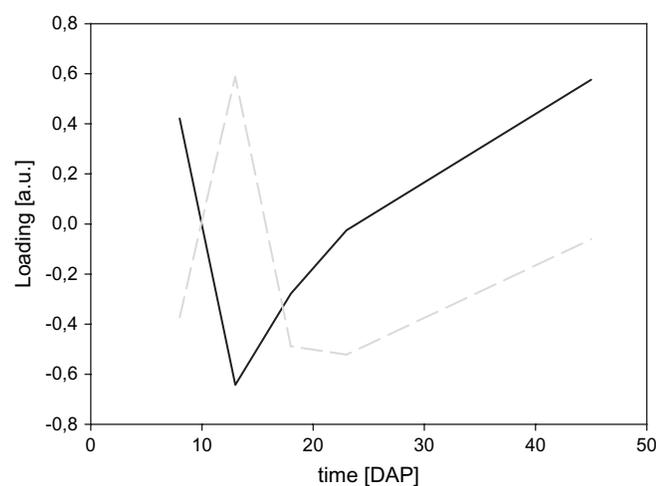

Fig. 8. Plot of the loading along the time dimension. The black full line refers to the first latent variable, while the grey dotted one to the second latent variable.

some cases the perturbation of the system can be detected observing that the factor loading of some metabolites have significantly different values, i.e they do not correlate with the same component in the control samples and in the perturbed ones. This variation can be 'concerted', underlying the invariant characteristics of the relation between metabolites that is resistant to the perturbation: groups of metabolites keep their correlation with the same component, while other do not keep it. In other cases, the signature of the perturbation can be revealed by a change of the sign of factor loading: metabolites positively correlated in the control samples have loading with the opposite sign in the perturbed ones (Camacho *et al.*, 2005; Manetti *et al.*, 2006).

When the metabonomic data set contains data collected at different times for different sample types, traditional

multivariate analysis increases the sample dimension of the unit/variable matrix but it 'distorts' the multivariate space spanned by the latent variables, mixing up the characteristics due to sample mode with those due to time mode. OPLS-DA tries to solve this 'problem' by introducing a dimension where the 'uninteresting' variation can be projected and treated as 'noise'.

In the N-way approach, time is explicitly considered and it generates more interpretable trajectories and is able clearly to represent the analysed samples whose characteristics have been known since the beginning.

It is important to underline that, in biological contexts, the number of samples is often limited and for this reason it is important not to forget any possible overfitting problems and not to stress the predictive capabilities of the models. For this reason, in this study, a non-supervised multi-way approach (PARAFAC) has also been used, in addition to the supervised one, confirming the results but at the same time showing that N-PLS-DA allows a better interpretation. In PARAFAC, no class information is presented to the model at the beginning, i.e. no Y matrix is defined.

Note that the estimated N-way model cannot be rotated without a loss of fit, as opposed to two-way analysis where one may rotate scores and loadings without changing the fit of the model: consequence of this characteristic (uniqueness of the solution) is more interpretable latent variables (Bro, 1997).

The peculiar dynamic in the early development of maize seeds comes out as being characteristic of the perturbation due to biological processes related to the different traits of the system (up- and down-regulation of *hda101*). In fact, putting together the results of score plot and loading plot analysis (Figs 6, 8), it is possible to observe that the factor distinguishing OE1-derived kernels from those of the other two genotypes has a time loading with different characteristics from the time loading of the second latent variable that discriminates between WT and modified kernels.

From the comparison of the two loadings, it is possible to say that the early stages of development are influenced by the introduction of up- and down-regulation of *hda101*. This can constitute a framework where results of univariate analysis can be collocated.

Our results indicated that kernels obtained from *hda101* over-expressing plants (OE1 line) exhibit a tendency to accumulate several metabolites compared with the levels observed in wild-type kernels (Tables 3, 4). On the contrary, kernels with down-regulation of *hda101* expression (AS33) show a tendency to a general decrease of metabolite levels. These different behaviours may be related to the involvement of *hda101* in controlling cell cycle progression, probably reflected in an alteration of the kernel size in *hda101* mutants in comparison to the wild type. In this respect, previous studies (Rossi *et al.*,

2003; Varotto *et al.*, 2003) have pointed out the involvement of *hda101* in cell cycle control, suggesting that hda101 physically interacts with the maize homologues of the mammalian retinoblastoma protein, a key regulator of G_1/S transition (Harbour and Dean, 2000). Therefore, it may be predicted that the up-regulation of HDA101 activity in the kernels of OE1 line induces aberrant effects on the normal progression of the cell cycle and, in particular, a general reduction of the cell number in developing seeds. This appears related to the decrease of the 100-kernel weight (Table 1) observed in the OE1 line compared to wild type (Rossi *et al.*, 2007) and to the profiles of the loadings relative to the time mode obtained by N-PLS-DA analysis (Fig. 8). The general repression of cell cycle progression may also result in a reduced capacity of the metabolites' utilization and transformation in the various metabolic pathways, leading to increased levels of metabolite accumulation observed in OE1 kernels. Conversely, the down-regulation of HDA101 activity in AS33 kernels may induce an increase in the rate of metabolites utilization to sustain the increased cell cycle activity. However, additional efforts are required to clarify this point. A similar general effect, determining an overall slow-down of early kernel development, was also observed by the over-expression of the activator ZmOCL1, a member of the ZmOCL (Outer Cell Layer) family, encoding putative transcription factors of the HD-ZIP IV class (Khaled *et al.*, 2005). In this context, the fact that HDA101 is usually considered as a repressor of gene transcription (Rossi *et al.*, 2007; Shahbazian and Grunstein, 2007), suggests that the opposite tendency of change in the metabolite levels in OE1 and AS33 kernels is due to opposite effects on the regulation of transcription of genes encoding for key regulators of metabolite synthesis and/or utilization.

Furthermore, these results reveal that the switch between the prestorage or cell division phase (8 DAP) to the storage or differentiation phase and mature seed is also accompanied by a switch from a much reduced hexose (α and β-glucose) and sucrose ratio to a high sucrose/hexose ratio within the embryo in the AS33 kernels to an increase of these metabolites in the OE1 kernels, in comparison to wild-type kernels (Tables 3, 4). It has been shown that a high hexose/sucrose ratio favours cell division whereas a high sucrose/hexose ratio favours differentiation to storage parenchyma cells. N-PLS-DA analysis also shows that sugars are key metabolites in discriminating among the three genotypes (Fig. 7). In fact, sugars such as glucose and sucrose can act as signals to trigger changes in gene expression in plants (Lam *et al.*, 1998) and post-transcriptional modification of proteins (Cotelle *et al.*, 2000) associated with nitrogen metabolism. These results have implicated a model in which genes involved in C and N metabolism are *cis*-regulated by both C and N signals (Coruzzi and Bush, 2001; Coruzzi and Zhou,

2001). Moreover, Price *et al.* (2004) found that glucose regulates a broader range of genes, incuding genes associated with carbohydrate metabolism, signal transduction, and metabolite transport. In addition, a large number of stress-responsive genes were also induced by glucose, indicating a role of sugar in environmental responses. It was also revealed that significant interactions exist between glucose and nitrogen in regulating gene expression, since glucose can modulate the effects of nitrogen and vice-versa.

A further observation which originates from the metabolic experiments was a significant variation in alanine content (Tables 3, 4; Fig. 7). This appears from our analyses one of the most important metabolites for discriminating among the three genopypes investigated here. In this context, it was shown that synthesis of alanine may occur at the expense of the amino acids, glutamate and aspartate (Stewart and Larher, 1980) and occurs concomitantly with the accumulation of 4-amino-butyrate or GABA (Ratcliffe, 1995), mediated by the enzyme alanine aminotransferase. This enzyme belongs to a pyridoxal phosphate multigene family widely distributed in animals, plants, algae, yeast, and bacteria (Vedavathi *et al.*, 2004). It catalyses transamination reactions using several amino donor:acceptor combinations, including the reversible transfer of an amino group from glutamate to pyruvate to form 2-oxoglutarate and alanine (Ricoult *et al.*, 2006). The observations are consistent with our finding and with that of Rossi *et al.* (2007) who reported from a functional analyses of the genotypes investigated here an appreciable alteration in expression level of the gene encoding alanine amino-transferase among the same genotypes investigated in this study.

In conclusion, this study shows results that clearly reveal the influence of *hda101* on maize metabolism, and at the same time it underlines the possibilities of multivariate analysis, used not to build a predictive model, but to describe the evolution of the system. Obviously, in this frame, multivariate analysis is a useful tool for physiologists, as they can synthesize experimental information in an interpretable form, that increases knowledge on the process studied in a way not possible by a traditional univariate approach.

Acknowledgements

References

Andersson CA, Bro R. 2000. The N-way toolbox for MATLAB. *Chemometrics and Intelligent Laboratory Systems* **52,** 1–4.

Antti H, Bollard ME, Ebbels T, Keun H, Lindon JC, Nicholson JK, Holmes E. 2002. Batch statistical processing of ^{1}H-NMR-derived urinary spectral data. *Journal of Chemometrics* **16,** 461–468.

Berger SL. 2007. The complex language of chromatin regulation during transcription. *Nature* **447,** 407–412.

Bosco MV, Garrido M, Larrechi MS. 2006. Determination of phenol in the presence of its principal degradation products in water during a TiO_2-photocatalytic degradation process by three-dimensional excitation–emission matrix fluorescence and parallel factor analysis. *Analytica Chimica Acta* **559,** 240–247.

Bro R. 1996. Multi-way calibration. Multilinear PLS. *Journal of Chemometrics* **10,** 47–61.

Bro R. 1997. PARAFAC. Tutorial and applications. *Chemometrics and Intelligent Laboratory Systems* **38,** 149–171.

Bylesjö M, Rantalainen M, Cloarec O, Nicholson JK, Holmes E, Trygg J. 2007. OPLS discriminant analysis: combining the strength of PLS-DA and SIMCA classification. *Journal of Chemometrics* **20,** 341–351.

Camacho D, de la Fuente A, Mendes P. 2005. The origin of correlations in metabolomics data. *Metabolomics* **1,** 53–63.

Carrozza MJ, Utley RT, Workman JL, Cote J. 2003. The diverse functions of histone acetyltransferase complexes. *Trends in Genetics* **19,** 321–329.

Castro C, Manetti C. 2007. A multi-way approach to analyze metabonomic data: a study of maize seeds development. *Analytical Biochemistry* **371,** 194–200.

Chow E, Ebrahimi D, Gooding JJ, Hibbert DB. 2006. Application of N-PLS calibration to the simultaneous determination of Cu^{2+}, Cd^{2+} and Pb^{2+} using peptide modified electrochemical sensors. *Analyst* **131,** 1051–1057.

Christensen AH, Sharrock RA, Quail PH. 1992. Maize poly-ubiquitin genes: structure, thermal perturbation of expression and transcript splicing, and promoter activity following transfer to protoplasts by electroporation. *Plant Molecular Biology* **18,** 675–689.

Christensen J, Nörgaard L, Bro R, Engelsen SB. 2006. Multivariate autofluorescence of intact food systems. *Chemistry Reviews* **106,** 1979–1994.

Cloarec O, Dumas ME, Craig A, et al. 2005b. Statistical total correlation spectroscopy: an exploratory approach for latent biomarker identification from metabolic ^{1}H NMR data sets. *Analytical Chemistry* **77,** 1282–1289.

Cloarec O, Dumas ME, Trygg J, Craig A, Barton RH, Lindon JC, Nicholson JK, Holmes E. 2005a. Evaluation of the orthogonal projection on latent structure model limitations caused by chemical shift variability and improved visualization of biomarker changes in ^{1}H-NMR spectroscopic metabonomic studies. *Analytical Chemistry* **77,** 517–526.

Coen M, Holmes E, Lindon JC, Nicholson JK. 2008. NMR-based metabolic profiling and metabonomic approaches to problems in molecular toxicology. *Chemical Research in Toxicology* **21,** 9–27.

Cotelle V, Meek SE, Provan F, Milne FC, Morrice N, MacKintosh C. 2000. 14–3–3s regulate global cleavage of their diverse binding partners in sugar-starved Arabidopsis cells. *EMBO Journal* **19,** 2869–2876.

Coruzzi G, Bush DR. 2001. Nitrogen and carbon nutrient and metabolite signaling in plants. *Plant Physiology* **125,** 61–64.

Coruzzi G, Zhou L. 2001. Carbon and nitrogen sensing and signaling in plants: emerging 'matrix effects'. *Current Opinion in Plant Biology* **4,** 247–253.

Durante C, Cocchi M, Grandi M, Marchetti A, Bro R. 2006. Application of N-PLS to gas chromatographic and sensory data of traditional balsamic vinegars of Modena. *Chemometrics and Intelligent Laboratory Systems* **83,** 54–65.

Dyrby M, Baunsgaard D, Bro R, Engelsen SB. 2005a. Multi-way chemometric analysis of the metabolic response to toxins monitored by NMR. *Chemometrics and Intelligent Laboratory Systems* **76,** 79–89.

Dyrby M, Peterson M, Whittaker AK, Lambert L, Nørgaard L, Bro R, Engelsen SB. 2005b. Analysis of lipoproteins using 2D diffusion-edited NMR spectroscopy and multi-way chemometrics. *Analytica Chimica Acta* **531,** 209–216.

Fan T. 1996. Metabolite profiling by one- and two-dimensional NMR analysis of complex mixtures. *Progress in Nuclear Magnetic Resonance Spectroscopy* **28,** 161–219.

Fischle W, Wang Y, Allis CD. 2003. Histone and chromatin cross-talk. *Current Opinion in Cell Biology* **15,** 172–183.

Guimet F, Ferreä J, Boqueä R, Vidal M, Garcia J. 2005. Excitation-emission fluorescence spectroscopy combined with three-way methods of analysis as a complementary technique for olive oil characterization. *Journal of Agricultural and Food Chemistry* **53,** 9319–9328.

Harbour JW, Dean DC. 2000. The Rb/E2F pathway: expanding roles and emerging paradigms. *Genes and Development* **14,** 2393–2409.

Jang IC, Pahk YM, Song SI, Kwon HJ, Nahm BH, Kim JK. 2003. Structure and expression of the rice class-I type histone deacetylase genes OsHDAC1–3: OsHDAC1 overexpression in transgenic plants leads to increased growth rate and altered architecture. *The Plant Journal* **33,** 531–541.

Jenuwein T, Allis CD. 2001. Translating the histone code. *Science* **293,** 1074–1080.

Kadonaga JT. 1998. Eukaryotic transcription: an interlaced network of transcription factors and chromatin-modifying machines. *Cell* **92,** 307–313.

Khaled AS, Vernoud V, Ingram GC, Perez P, Sarda X, Rogowsky PM. 2005. Engrailed-ZmOCL1 fusions cause a transient reduction of kernel size in maize. *Plant Molecular Biology* **58,** 123–139.

Keun HC, Ebbels TMD, Bollard ME, Beckonert O, Antti H, Holmes E, Lindon JC, Nicholson JK. 2004. Geometric Trajectory Analysis of metabolic responses to toxicity can define treatment specific profiles. *Chemical Research in Toxicology* **17,** 579–587.

Lam HM, Chiu J, Hsieh MH, Meisel L, Oliveira IC, Shin M, Coruzzi G. 1998. Glutamate-receptor genes in plants. *Nature* **396,** 125–126.

Le Gall G, Colquhoun IJ, Davis AL, Collins GJ, Verhoeyen ME. 2003. Metabolite profiling of tomato *(Lycopersicon esculentum)* using ^{1}H NMR spectroscopy as a tool to detect potential unintended effects following a genetic modification. *Journal of Agricultural and Food Chemistry* **51,** 2447–2456.

Lindon JC, Holmes E, Nicholson JK. 2001. Pattern recognition methods and application in biomedical magnetic resonance. *Progress in Nuclear Magnetic Resonance Spectroscopy* **39,** 1–40.

Loidl P. 2004. A plant dialect of the histone language. *Trends in Plant Science* **9,** 84–90.

Long JA, Ohno C, Smith ZR, Meyerowitz EM. 2006. TOPLESS regulates apical embryonic fate in *Arabidopsis*. *Science* **312,** 1520–1523.

Lusser A, Kölle D, Loidl P. 2001. Histone acetylation: lessons from plant kingdom. *Trends in Plant Science* **6,** 59–65.

Manetti C, Bianchetti C, Bizzarri M, et al. 2004. NMR-based metabonomics study of transgenic maize. *Phytochemistry* **65,** 3187–3198.

Manetti C, Bianchetti C, Casciani L, Castro C, Di Cocco ME, Miccheli A, Motto M, Conti F. 2006. A metabonomic study of transgenic maize (*Zea mays*) seeds revealed variation in osmolytes and branched amino acids. *Journal of Experimental Biology* **57,** 2613–2625.

Miccheli A, Aureli T, Delfini M, Di Cocco ME, Viola P, Gobetto R, Conti F. 1988. Study on influence of inactivation enzyme techniques and extraction procedures on cerebral phosphorylated metabolite levels by P-31 NMR spectroscopy. *Cellular and Molecular Biology* **34,** 591–603.

Morocz S, Donn G, Nemeth J, Dudits D. 1990. An improved system to obtain fertile regenerants via maize protoplasts isolated from a highly embryogenic suspension culture. *Theoretical and Applied Genetics* **80,** 721–726.

Niazi A, Ghasemi J, Yazdanipour A. 2005. PARAFAC decomposition of three-way kinetic-spectrophotometric spectral matrices based on phosphomolybdenum blue complex chemistry for nitrite determination in water and meat samples. *Analytical Letters* **38,** 2377–2392.

Nicholson JK, Connelly J, Lindon JC, Holmes E. 2002. Metabonomics: a platform for studying drug toxicity and gene function. *Nature Reviews Drug Discovery* **1,** 153–161.

Pandey R, Muller A, Napoli CA, Selinger DA, Pikaard CS, Richards EJ, Bender J, Mount DW, Jorgensen RA. 2002. Analysis of histone acetyltransferase and histone deacetylase families of *Arabidopsis thaliana* suggests functional diversification of chromatin modification among multicellular eukaryotes. *Nucleic Acids Research* **30,** 5036–5055.

Peterson CL, Laniel M. 2004. Histones and histone modifications. *Current Biology* **14,** R546–R551.

Pfluger J, Wagner D. 2007. Histone modifications and dynamic regulation of genome accessibility in plants. *Current Opinion in Plant Biology* **10,** 645–652.

Povlsen VT, Rinnan Å, van den Berg F, Andersen HJ, Thybo AK. 2003. Direct decomposition of NMR relaxation profiles and prediction of sensory attributes of potato samples. *LWT-Food Science and Technology* **36,** 423–432.

Price J, Laxmi A, St Martin SK, Jang JC. 2004. Global transcription profiling reveals multiple sugar signal transduction mechanisms in *Arabidopsis*. *The Plant Cell* **16,** 2128–2150.

Probst AV, Fagard M, Proux F, *et al.* 2004. Arabidopsis histone deacetylase HDA6 is required for maintenance of transcriptional gene silencing and determines nuclear organization of rDNA repeats. *The Plant Cell* **16,** 1021–1034.

Ricciolini R, Miccheli A, Di Cocco ME, Piccolella E, Marino A, Sammartino MP, Conti F. 1994. Dexamethasone-dependent modulation of human lymphoblastoid B cell line through sphingosine production. *Biochimica et Biophysica Acta* **1221,** 103–108.

Ratcliffe RG. 1995. Metabolic aspects of the anoxic response in plant tissue. In: Smirnoff N, ed. *Environment and plant metabolism: flexibily and acclimatation*. Oxford, UK: Bios Scientific, 111–189.

Ricoult C, Echeverria HO, Cliquet J-B, Limani AM. 2006. Charactecterization of alanine aminotransferase (*AlaAT*) multigene family and hypoxic response in young seedlinds of the model legume Medicago truncatula. *Journal of Experimental Biology* **37,** 3079–3089.

Rossi V, Hartings H, Motto M. 1998. Identification and characterisation of an RPD3 homologue from maize (*Zea mays* L.) that is able to complement an *rpd3* null mutant of *Saccharomyces cerevisiae*. *Molecular and General Genetics* **258,** 288–296.

Rossi V, Locatelli S, Lanzanova C, *et al.* 2003. A maize histone deacetylase and retinoblastoma-related protein physically interact and cooperate in repressing gene transcription. *Plant Molecular Biology* **51,** 401–413.

Rossi V, Locatelli S, Varotto S, Donn G, Pirona R, Henderson DA, Hartings H, Motto M. 2007. Maize histone deacetylase hda101 is involved in plant development, gene transcription, and sequence-specific modulation of histone modification of genes and repeats. *The Plant Cell* **19,** 1145–1162.

Shahbazian MD, Grunstein M. 2007. Functions of site-specific histone acetylation and deacetylation. *Annual Review of Biochemistry* **76,** 1–26.

Stewart GR, Larher F. 1980. Accumulation of amino acids and related compounds in relation to environmental stress. In: Miflin BJ, ed. *The biochemistry of plants*, Vol 5. New York, NY: Academic Press, 609–635.

Thiagalingam S, Cheng KH, Lee HJ, Mineva N, Thiagalingam A, Ponte JF. 2003. Histone deacetylases: unique players in shaping the epigenetic histone code. *Annals of the New York Academy of Science* **983,** 84–100.

Tian L, Chen ZJ. 2001. Blocking histone deacetylation in Arabidopsis induces pleiotropic effects on plant gene regulation and development. *Proceedings of the National Academy of Sciences, USA* **98,** 200–205.

Trygg J, Wold S. 2002. Orthogonal projection to latent structure. *Journal of Chemometrics* **16,** 283–293.

Varotto S, Locatelli S, Canova S, Pipal A, Motto M, Rossi V. 2003. Expression profile and cellular localization of maize Rpd3-type histone deacetylase during plant development. *Plant Physiology* **133,** 606–617.

Vedavathi M, Girish KS, Kumar MK. 2004. Isolation and characterization of cytosolic alanine aminotransferase isoforms from starved rat liver. *Molecular and Cellular Biochemistry* **267,** 13–23.

Zhou C, Zhang L, Duan J, Miki B, Wu K. 2005. *Histone Deacetylase19* is involved in jasmonic acid and ethylene signaling of pathogen response in *Arabidopsis*. *The Plant Cell* **17,** 1196–1204.