

A metabolomic study of substantial equivalence of field-grown genetically modified wheat

John M. Baker, Nathaniel D. Hawkins, Jane L. Ward, Alison Lovegrove, Johnathan A. Napier, Peter R. Shewry and Michael H. Beale*

National Centre for Plant and Microbial Metabolomics, Crop Performance and Improvement Division, Rothamsted Research, Harpenden, Hertfordshire AL5 2JQ, UK

Received 20 January 2006;

revised 5 April 2006;

accepted 11 April 2006

*Correspondence (fax +44 1582 763010;

e-mail mike.beale@bbsrc.ac.uk)

Summary

The 'substantial equivalence' of three transgenic wheats expressing additional high-molecular-weight subunit genes and the corresponding parental lines (two lines plus a null transformant) was examined using metabolite profiling of samples grown in replicate field trials on two UK sites (Rothamsted, Hertfordshire and Long Ashton, near Bristol) for 3 years. Multivariate comparison of the proton nuclear magnetic resonance spectra of polar metabolites extracted with deuterated methanol–water showed a stronger influence of site and year than of genotype. Nevertheless, some separation between the transgenic and parental lines was observed, notably between the transgenic line B73-6-1 (which had the highest level of transgene expression) and its parental line L88-6. Comparison of the spectra showed that this separation resulted from increased levels of maltose and/or sucrose in this transgenic line, and that differences in free amino acids were also apparent. More detailed studies of the amino acid composition of material grown in 2000 were carried out using gas chromatography-mass spectrometry. The most noticeable difference was that the samples grown at Rothamsted consistently contained larger amounts of acidic amino acids (glutamic, aspartic) and their amides (glutamine, asparagine). In addition, the related lines, L88-6 and B73-6-1, both contained larger amounts of proline and γ -aminobutyric acid when grown at Long Ashton than at Rothamsted. The results clearly demonstrate that the environment affects the metabolome and that any differences between the control and transgenic lines are generally within the same range as the differences observed between the control lines grown on different sites and in different years.

Keywords: gas chromatography-mass spectrometry (GC-MS), genetically modified (GM) crops, high-molecular-weight (HMW) subunits, metabolic fingerprinting, metabolomics, nuclear magnetic resonance (NMR), substantial equivalence, wheat.

Introduction

The concerns of consumers and regulatory authorities mean that transgenic crops are currently being subjected to detailed scrutiny to assess their safety for use in food. A major point of concern is that it is not currently possible to target the transgenes to specific sites within the genome of the host plant, and that little is known about the mechanisms that determine the sites at which integration occurs. Furthermore, plants produced by biolistics may contain multiple transgene inserts, including rearranged and partial copies, and at several loci (Barcelo *et al.*, 2001; Rooke *et al.*, 2003). This random insertion has led to the suggestion that transgenes may have

unpredicted effects on the expression of other endogenous genes, depending on their sites of insertion. Although there is currently little evidence to support this (see recent reviews by Kohli *et al.*, 2003; Sahrawat *et al.*, 2003; Altpeter *et al.*, 2005; Jones, 2005), even an infrequent event could have serious consequences if it resulted in the production of deleterious components or the loss of components which currently contribute to nutritional or processing quality. The principal criterion that has been put forward for assessing the safety of novel foodstuffs is that of substantial equivalence (OECD, 1993; FAO/WHO, 2000; Kuiper *et al.*, 2001; Kok and Kuiper, 2003), the principle behind this being that, if the composition of the novel (i.e. transgenic) foodstuff does not differ in a

meaningful way from a traditional (i.e. unmodified) variety(s), it is safe. It is important to take into account the 'natural variability' in the composition of the traditional variety when making this comparison. In the case of major crops, this 'natural variability' may be a result of a diverse range of effects (weather, soil, location, etc.), whose influences are impossible to predict.

We have carried out a series of metabolomic studies to address these concerns using field-grown transgenic wheat lines developed at Rothamsted Research. These lines contain additional copies of genes encoding a group of endogenous wheat gluten proteins called the high-molecular-weight (HMW) subunits of glutenin. These HMW subunits are only synthesized in the starchy endosperm cells of the developing wheat grain, accounting for up to approximately 12% of the total proteins at maturity (Shewry *et al.*, 2003). They have no known biological function except storage, and hence manipulation of their amount and composition would not be expected to have a major effect on the synthesis of other grain components. However, for food processing applications, HMW subunits of glutenin have important functional properties, with observed allelic variation in their composition being associated with differences in the elasticity of gluten and dough (Payne, 1987; Shewry *et al.*, 2003). Hence, specific alleles are routinely selected by plant breeders (Payne, 1987), and such alleles were an early target for attempts to improve wheat processing quality by genetic engineering (Altpeter *et al.*, 1996; Blechl and Anderson, 1996; Barro *et al.*, 1997; Alvarez *et al.*, 2000). Transgenic lines expressing additional HMW subunit genes therefore provide an excellent example of the type of manipulation which may be applied to improve wheat processing quality in the future, and represent ideal material to compare the 'substantial equivalence' of genetically modified (GM) and conventionally bred crops.

We have grown a series of transgenic wheat lines expressing additional HMW subunit genes and appropriate control lines in replicate field experiments on two sites in the UK for 4 years (1998–2001). Material from these experiments has been used to establish that the transgenes do not differ from the endogenous HMW subunit genes in their stability of expression, and that the transgenic and control lines are similarly stable when comparing grain weight, grain nitrogen content and dough mixing properties (Shewry *et al.*, 2006). An accompanying paper (Baudo *et al.*, 2006) reports detailed transcriptome profiling of selected lines from the same series. However, we did not establish the 'substantial equivalence' of these lines at the detailed compositional level, and this is addressed here using a hierarchical metabolomics approach employing proton nuclear magnetic resonance (^1H NMR) metabolite fingerprinting combined with more detailed gas

chromatography-mass spectrometry (GC-MS) analyses of selected samples. ^1H NMR has been successfully used in plant metabolomics studies (Ward *et al.*, 2003) and in the assessment of the substantial equivalence of GM and control lines in several studies (Le Gall *et al.*, 2003; Charlton *et al.*, 2004; Choi *et al.*, 2004; Defernez *et al.*, 2004; Manetti *et al.*, 2004; Garratt *et al.*, 2005; McCann *et al.*, 2005; Oberdoerfer *et al.*, 2005), but this is the first example of its application to field-grown material of a major crop which has been engineered to improve a quality trait.

Results and discussion

Characteristics of the lines

The characteristics of the lines used are summarized in Table 1. The two parental lines, L88-6 and L88-31, are both derived from T_2 seed of a cross between the Australian cultivars Olympic and Gabo, and both have 50% background from each parent. Hence, they are related but not isogenic. The presence of null alleles for HMW subunit genes in the parental lines allowed the selection of a series of lines with 0–5 expressed HMW subunit genes, and L88-6 and L88-31 form part of this series (Lawrence *et al.*, 1988). The transgenic lines B102-1-1, B102-1-2 and B73-6-1 were produced by biolistic transformation with two plasmids: one carried the gene of interest (encoding HMW subunit 1Ax1 or 1Dx5) under the control of its own starchy endosperm-specific promoter, and the second was pACH 25 containing the *bar* and *uidA* genes under the control of the maize ubiquitin promoter (Christensen and Quail, 1996). Finally, line B72-8-11a was a null segregant from the transformation of L88-31 and was included as an additional control. Thus, with the exception of B102-1-2 (which appears to have lost the *uidA* gene by plasmid rearrangement) (Rooke *et al.*, 2003), all six transgenic lines should also express the *bar* and *uidA* genes in all tissues.

These lines were grown in replicate field trials on two sites [Long Ashton Research Station (LARS) and Rothamsted Research (RRes)] over 4 years (1998–2001), and plots from 1999 to 2001 (corresponding to T_5 , T_6 and T_7 grain) were selected for analysis. Four randomly selected plots from each line/year/site combination were milled to give white flour. White flour was selected for analysis for several reasons. Firstly, this corresponds to the starchy endosperm tissue in which the HMW subunit transgenes are expressed; secondly, it is the part of the grain most widely used in food products; and, thirdly, it avoids the effect of seed size, which could influence the results of such wholemeal analyses (by affecting the proportions of bran, germ and flour components in the meal).

Table 1 Characteristics of the control and transgenic lines used in this study

Line	Characteristics	Endogenous HMW subunit genes	HMW subunit transgenes		Marker genes
			Gene	Insertions	
L88-6	Parental line, derived from Olympic × Gabo	1Ax1, 1Bx17, 1By18, 1Dx5, 1Dy10	None	–	None
B73-6-1	L88-6 transformed with 1Dx5 gene	1Ax1, 1Bx17, 1By18, 1Dx5, 1Dy10	1Dx5	3 loci 10–15 copies	<i>bar</i> , <i>uidA</i>
L88-31	Parental line. Sister line derived from same cross as L88-6	1A null, 1Bx17, 1By18, 1D null	None	–	None
B102-1-1	L88-31 transformed with Ax1 gene	1A null, 1Bx17, 1By18, 1D null	1Ax1	1 locus 2–4 copies	<i>bar</i> , <i>uidA</i>
B102-1-2	L88-31 transformed with 1Ax1 gene. Sister line of B102-1-1	1A null, 1Bx17, 1By18, 1D null	1Ax1	2 loci 4–5 copies	<i>bar</i> *
B72-8-11a	Null transformant of L88-31	1A null, 1Bx17, 1By18, 1D null	None	–	None

HMW, high molecular weight.

The table is based on data in Lawrence *et al.* (1988), Barro *et al.* (1997) and Rooke *et al.* (2003).

**uidA* not detected by polymerase chain reaction; presumably lost by plasmid rearrangement during transformation.

Metabolite fingerprinting analysis by ¹H NMR

The NMR samples were prepared and analysed in three separate groups, each group containing all of the flours from one of the 3 years of field trials. A typical ¹H NMR spectrum of a polar extract from a white flour sample is shown in Figure 1. The spectrum is dominated by overlapping carbohydrate peaks (between δ3.5 and 4.0) (Figure 1a), with smaller peaks for the anomeric protons of sucrose, maltose and glucose between δ4.0 and 5.5 (Figure 1b). In addition, signals from amino acids can clearly be seen in the aliphatic region (Figure 1c). Similar spectra were collected from samples prepared and analysed in batches of 11 per day, with the data being recorded less than 48 h after extract preparation. The order and day on which samples were run were determined using a randomized block experimental design. In order to ensure that samples run on different days or in different groups were comparable, a 12th sample of control flour was extracted and analysed each day. The flour from wheat grown in plot 16 at RRes in 2000 (line L88-6) was used as this control. When principal component analysis (PCA) was performed on all of the data, the control samples formed a tight cluster in the first four components (accounting for 92% of sample variance) (Figure 2). This indicated that all of the samples were comparable and that no variation as a result of temporal experimental error and/or instrument drift had been introduced. For all subsequent data analysis, these tracking control sample data were removed.

The NMR spectra, collected from three replicate extracts of each of the flour samples (six lines × two sites × three years × four plots), were compared by PCA (scores plot; Figure 3).

The first two principal components accounted for 73% and 15% of the variance, respectively. The model demonstrates that there is a considerable environmental effect on the metabolome, with individual year–site groups forming separate clusters. The broad appearance of the year–site clusters (and, in the case of RRes 1999, separation into two clusters) can be attributed to differences between individual lines. The major differences across the individual year–site clusters are highlighted in Figure 4. This displays the same PCA scores model, with individual year–site data sets highlighted in turn. As can be seen, there is generally a separation between the transgenic line B73-6-1 and the other GM and parental lines. This is most clear in RRes 1999-grown material, where B73-6-1 is clearly well separated from all other (GM and non-GM) lines. The other three site–year combinations from 1999 and 2000 show a similar (if smaller) effect. The fact that separations between B73-6-1 and the other lines can be seen in models of the total data set clearly indicates that such effects are significant when compared with the natural variability (i.e. differences due to year, site and plot).

In order to better examine the differences between B73-6-1 and its parent line L88-6, PCA models for each site–year were constructed, using only these lines (Figure 5). Again these models showed a separation, in principal component 1 (PC1), between the L88-6 and B73-6-1 samples grown at both sites in 1999 and 2000. The loadings plots for PC1 describe the differences in metabolite levels between these two lines and account for their separation in the scores plot. Analysis of the loadings plots of PC1 for three of these models (LARS 1999, LARS 2000 and RRes 1999) showed increased levels of

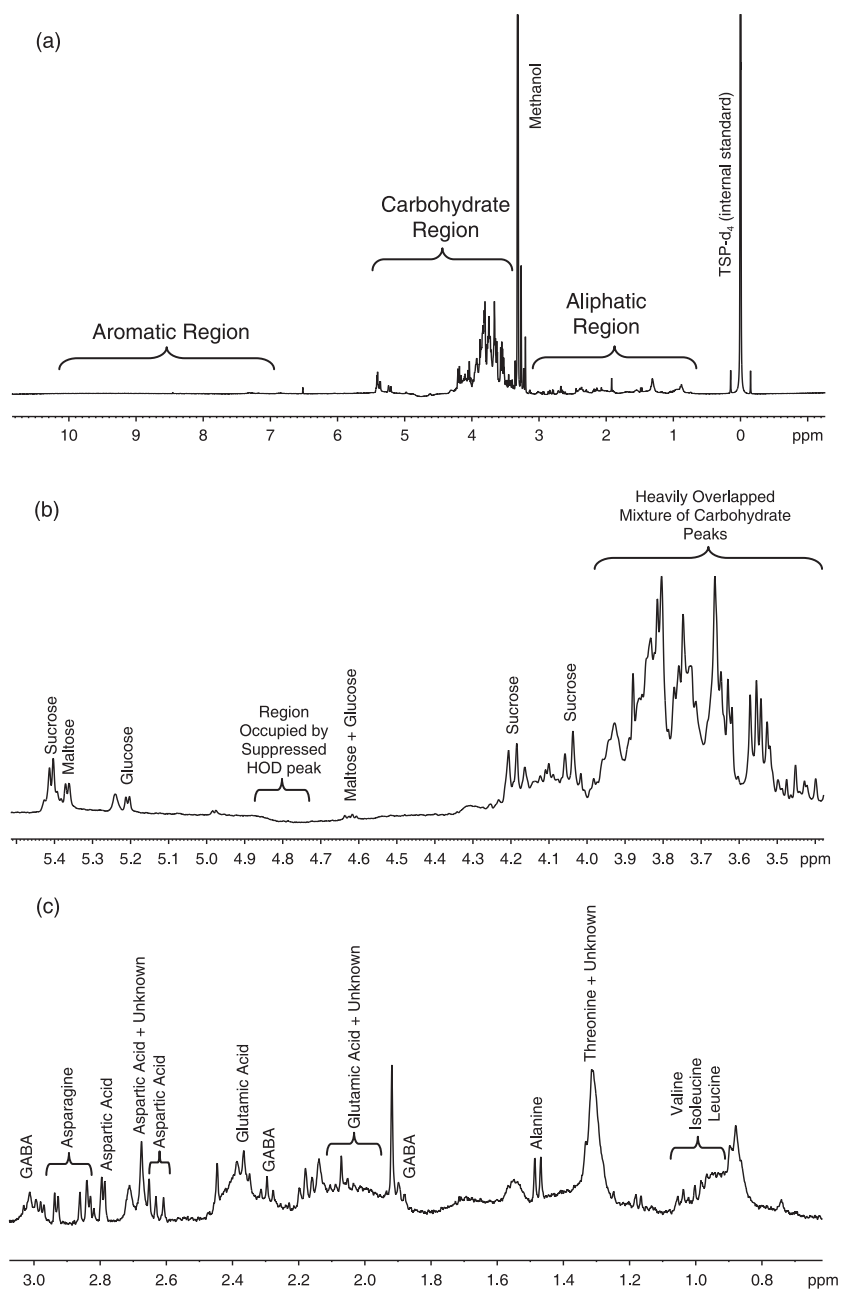


Figure 1 Proton nuclear magnetic resonance (^1H NMR) spectrum of an 80–20 D_2O – CD_3OD extract of white flour from grain of line L88-6 grown at Rothamsted in 2000: (a) whole spectrum; (b) expansion of carbohydrate region; (c) expansion of aliphatic region. GABA, γ -aminobutyric acid.

maltose in the transgenic line, as illustrated in Figure 5d for LARS 2000. In the fourth model (RRes 2000), PC1 represented increased levels of sucrose in the transgenic line. The models of L88-6 and B73-6-1 samples grown at either site in 2001 did not show such clustering.

The PCA model of the total data set was dominated by the differences between B73-6-1 and its parent line, L88-6. In order to identify any effects of the transgenes on the L88-31 family of lines (L88-31, B102-1-1, B102-1-2 and B72-8-11a), these lines were modelled separately. The scores plot for the first two PCs (accounting for 73.4% and 14.4% of the variance, respectively) is shown in Figure 6. Each site–year set

forms a single cluster with no discernible clustering by line. However, different field plots of the same line form clearly separated clusters (see Figure 6d). This shows that any effects on the metabolite profile resulting from the genetic modification of L88-31 are of lower magnitude than the effects of the plot in which the grain was grown. When the higher PCs of the model (PC3–PC6; accounting for 6.8% of the total variance between them) were examined, no clustering by line could be observed (data not shown).

It can be concluded from these analyses that the site, year and even the individual plot have a stronger influence on the metabolite profiles than does the genotype, for the lines

Figure 2 Principal components analysis (PCA) scores plots of all nuclear magnetic resonance (NMR) spectra, illustrating the use of the quality control tracking sample: red ●, tracking sample spectra; ▲, all other spectra.

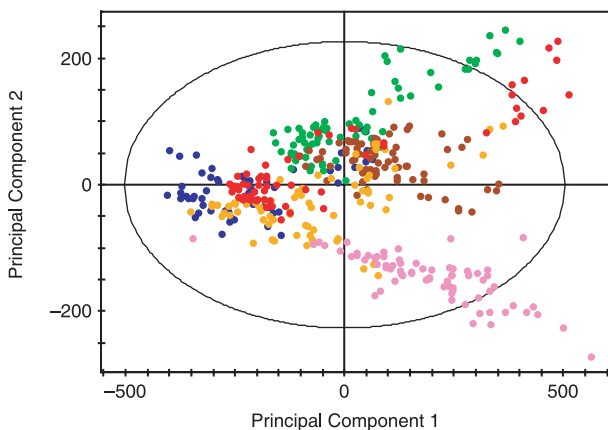
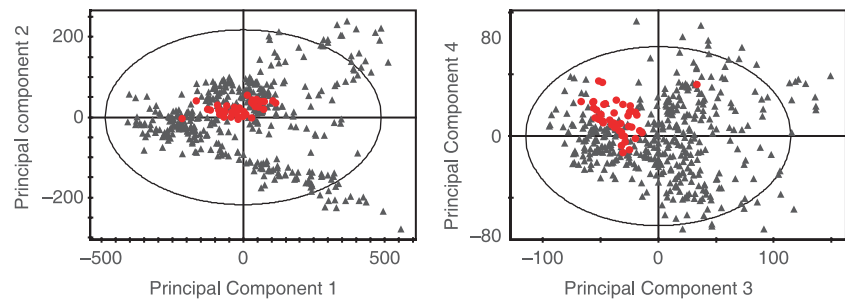


Figure 3 Principal components analysis (PCA) scores plot for PC1 and PC2 generated from data from all samples and coloured according to year and site: colour?, RRes 2001; brown ●, RRes 2000; red ●, RRes 1999; orange ●, LARS 2001; green ●, LARS 2000; blue ●, LARS 1999. LARS, Long Ashton Research Station; RRes, Rothamsted Research.

derived from L88-31. However, some separation between the parental and transgenic lines can be observed with L88-6 and B73-6-1. The differences observed between these lines were mainly due to changes in levels of the disaccharides maltose and sucrose. Differences between sites, determined from an analysis of the loadings plots (e.g. year 2000; Figure 7b) generated from PCA models of individual year data sets (Figure 7a), can be attributed in part to differences in levels of amino acids, particularly aspartate and glutamine. NMR spectroscopy was clearly ideal for the rapid comparison of multiple samples, but the spectra only provided quantitative information on differences in the more abundant polar analytes. To further investigate and quantify the differences in amino acid levels by site, one year set (2000) was analysed by a targeted GC-MS screen.

Amino acid analysis by GC-MS

The commercial kit used concentrated the basic components, present in a flour extract, on to a cation exchange resin. This basic fraction was then derivatized prior to GC-MS. Norvaline

was used as an internal standard and calibration curves were constructed for each amino acid to be measured. From these curves, the analyte concentrations for each of the samples were calculated from the relative peak areas of the relevant analytes. The most relevant results are displayed in Figure 8 as bar charts showing the concentration of each amino acid for each of the 48 flours investigated. The values used are the mean concentrations from the five replicate analyses of each flour sample, with the error bars representing the 95% confidence intervals for each concentration. The confidence intervals of the results for methionine, β -alanine and ornithine, compounds that were only present in trace amounts, were sufficiently large to prohibit any attempt to interpret them directly.

Analysis of variance (ANOVA) was used to determine the least significant difference values for the concentration of each amino acid between samples. The most obvious significant differences were between the flours grown at the two sites, with the samples grown at RRes consistently exhibiting higher concentrations of asparagine (c. 170 $\mu\text{g/g}$ difference) than those grown at LARS. Additional significant differences observed by GC-MS were elevated levels of aspartic acid (c. 250 $\mu\text{g/g}$ difference), glutamine (c. 90 $\mu\text{g/g}$ difference) and glutamic acid (c. 70 $\mu\text{g/g}$ difference) in the RRes-grown samples compared with those grown at LARS. L88-6 and B73-6-1 were equivalent in amino acid composition, but both showed an environmental effect in that they contained greater amounts of proline and γ -aminobutyric acid when grown at LARS but not at RRes. This effect was not seen with the lines derived from L88-31. The biological reason for this is not known. No evidence of any significant differences between the amino acid compositions of the control and transgenic lines was found.

General discussion

The analyses presented here represent one of the most detailed studies of the substantial equivalence of field-grown GM and non-GM crops at the metabolite level reported

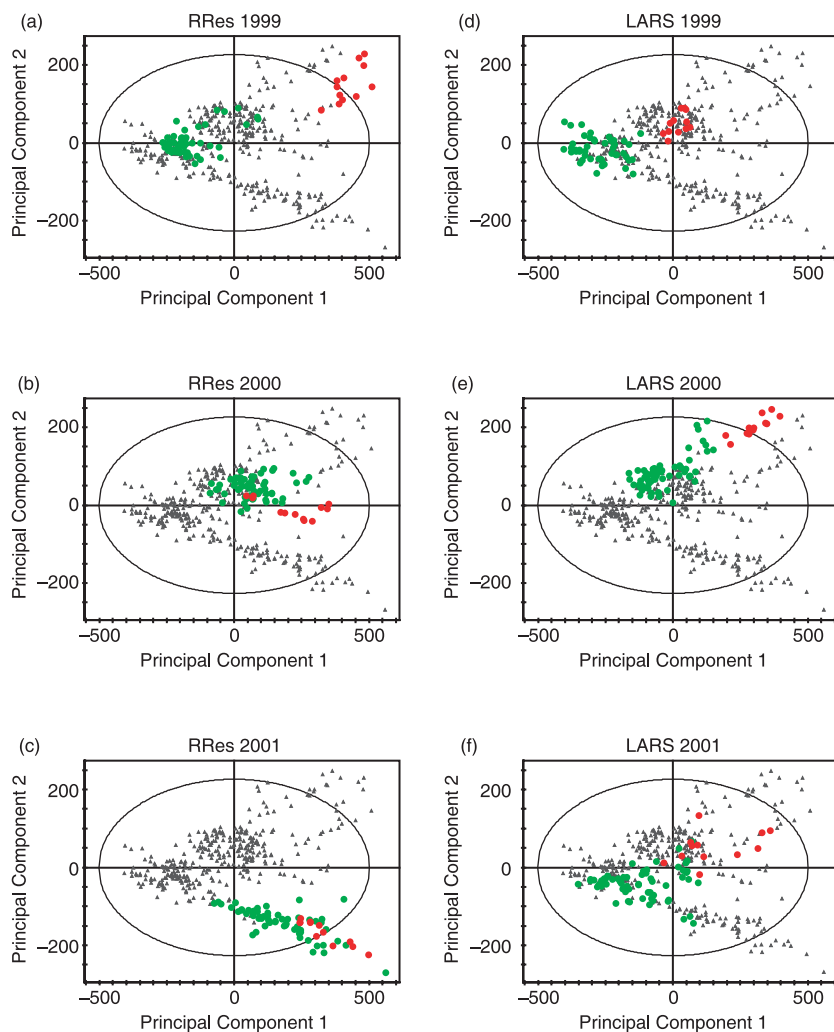


Figure 4 Principal components analysis (PCA) scores plot for PC1 and PC2 from the PCA model of the total nuclear magnetic resonance (NMR) data set coloured in six different ways; each plot highlights the clustering of B73-6-1 (red ●) and all other lines (green ●) grown at: (a) RRes in 1999; (b) RRes in 2000; (c) RRes in 2001; (d) LARS in 1999; (e) LARS in 2000; (f) LARS in 2001. LARS, Long Ashton Research Station; RRes, Rothamsted Research.

so far. Previous studies have analysed plant material grown under controlled conditions. These cannot reflect the 'natural variability' in the plants resulting from variation in environmental conditions of field trials conducted over several years and at different sites. Furthermore, the lines were selected to represent the type of modification that may in the future be carried out to improve grain quality, and the transgenes would be expected to have a minimal impact on grain metabolism, as their role is as storage proteins which are digested during seed germination.

Two technologies were applied to the metabolite screen to generate complementary data. NMR profiling provided a rapid, unbiased and reliable method for the assessment of differences in all abundant, proton-containing soluble metabolites. Differences between the amounts of carbohydrates and some amino acids in the wheat lines were readily visible by this technique. Furthermore, GC-MS of propyloxy-carbonylated derivatives proved to be a reliable technique

for the quantification of most of the amino acids. The use of broad brush metabolite fingerprinting, followed by targeted analysis of metabolites highlighted in the first screen, proved to be a productive strategy in the study of such field-grown experimental material.

Comparison of the null transformant (B72-8-11a) with the control line (L88-31), and of the transgenic lines with their respective controls, demonstrates that neither the transformation process nor the expression of the transgenes had significant and reproducible effects on the grain metabolome. Indeed, greater differences were often present between the two parental lines (L88-6, L88-31) and between the material from different sites and years than between the transgenic and control lines. However, differences were observed between lines L88-6 and B73-6-1 in four of the six data sets. The transgenic line B73-6-1 was derived from L88-6 by the addition of 10–15 copies of the subunit 1Dx5 transgene, and the level of transgene expression was high, with the proportion

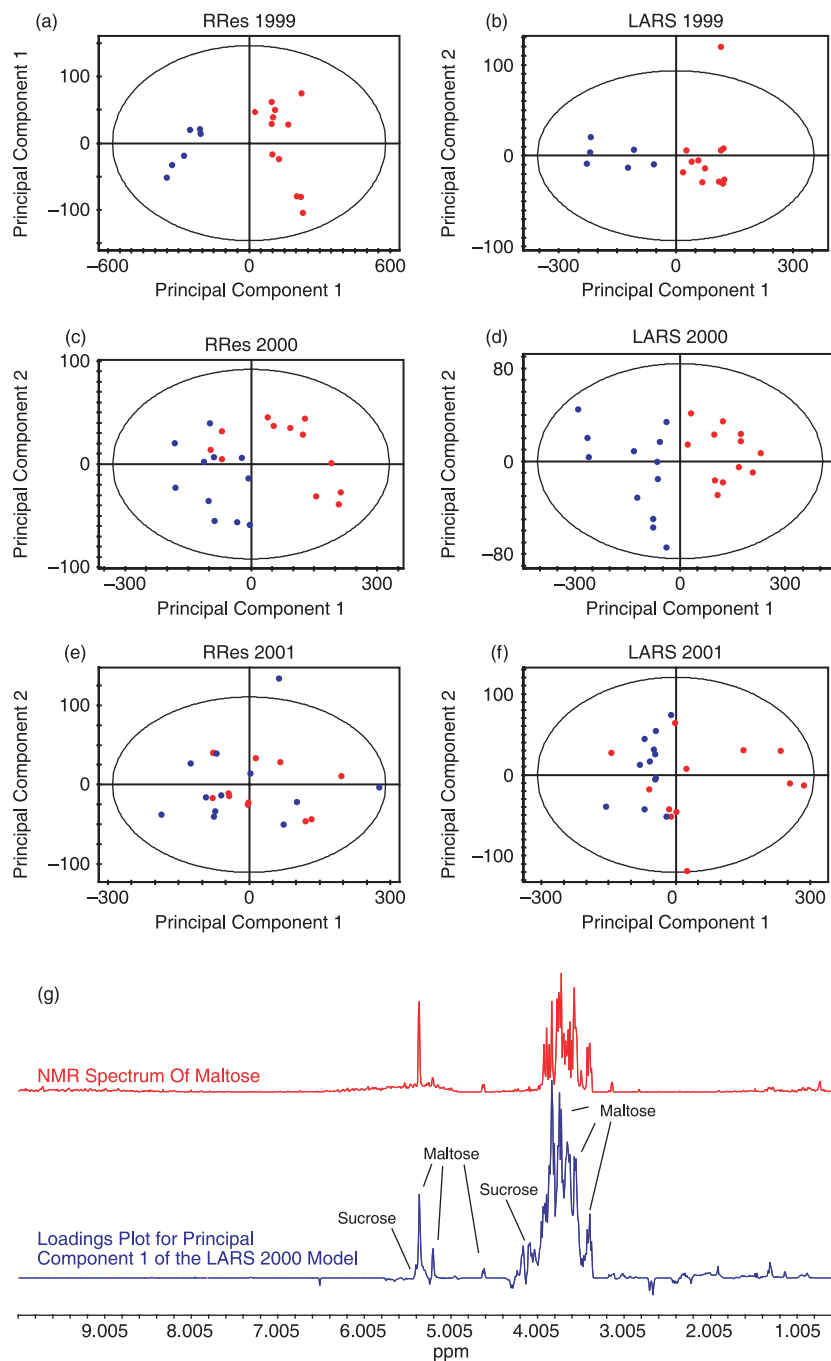


Figure 5 Principal components analysis (PCA) modelling of nuclear magnetic resonance (NMR) data from lines L88-6 (red ●) and B73-6-1 (blue ●). Scores plots from samples grown at: (a) RRes in 1999; (b) LARS in 1999; (c) RRes in 2000; (d) LARS in 2000; (e) RRes in 2001; (f) LARS in 2001. (g) Loadings plot for PC1 from the LARS 2000 model indicating that maltose is the main contributing metabolite. LARS, Long Ashton Research Station; RRes, Rothamsted Research.

of subunit 1Dx5 increasing from 3%–4% of the total grain protein in L88-6 to 10%–12% in B73-6-1 (Rooke *et al.*, 1999; Shewry *et al.*, 2006). This contrasts with B102-1-1 and B102-1-2, in which the numbers of transgene copies were lower (Table 1), and the proportions of the total protein encoded by the transgenes were lower (typically 4%–6%, depending on the plots) (Shewry *et al.*, 2006). Hence, it is not surprising that the detailed metabolomic study reported here showed some differences between L88-6 and B73-6-1, but little difference

between the other lines. However, it should be noted that even the differences between B73-6-1 and L88-6 were not reproducible across all site-year combinations, and that differences between sites were generally greater than differences between lines. Furthermore, the components on which these discriminations were made were common abundant metabolites.

The results of this study are consistent with those reported by Baudo *et al.* (2006) on transcriptome profiling of the

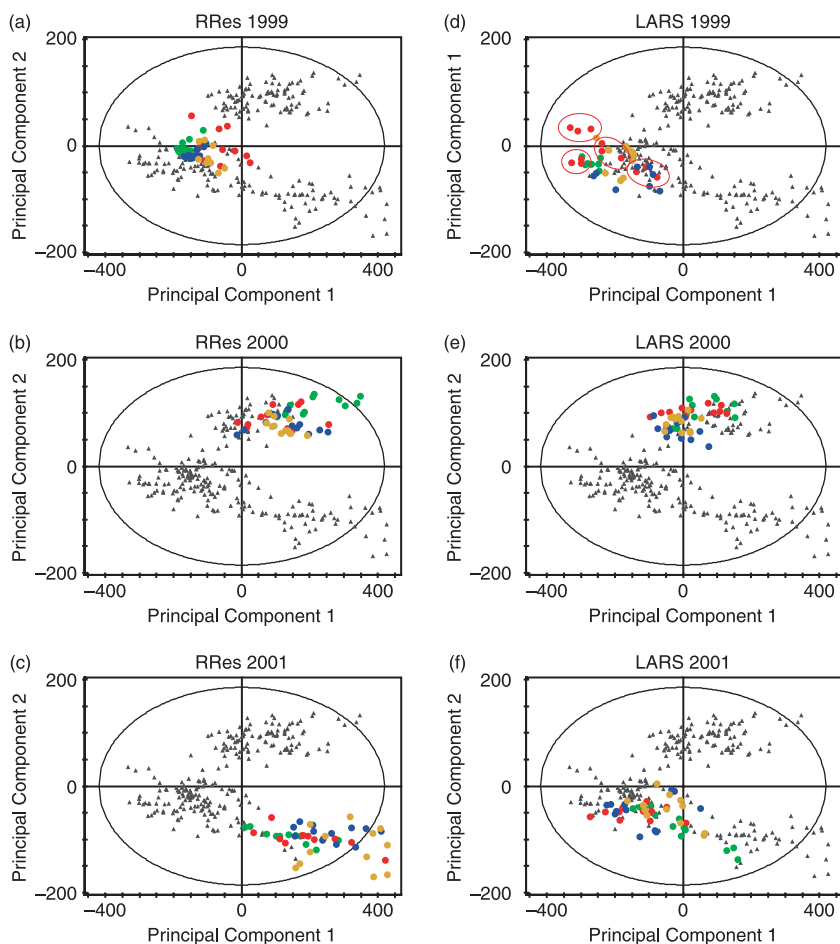


Figure 6 Scores plot for PC1 and PC2 from the principal components analysis (PCA) model of the L88-31 (orange ●), B102-1-1 (blue ●), B102-1-2 (red ●) and B72-8-11 (green ●) lines alone. Individual panels highlight different year–site sets. (a) RRes in 1999; (b) RRes in 2000; (c) RRes in 2001; (d) LARS in 1999; (e) LARS in 2000; (f) LARS in 2001. (d) Also shows the clustering of individual plots of B102-1-2 grown at LARS in 1999. LARS, Long Ashton Research Station; RRes, Rothamsted Research.

same series of lines. They used cDNA arrays to compare the transcriptome profiles of developing grain and leaves of L88-31, L88-18 (a sister line related to L88-31, but also expressing the endogenous 1Ax1 gene) and B102-1-1 (expressing the 1Ax1 transgene), and found no significant effects of the transgene on the global pattern of gene expression. Similarly, no significant differences were found between transgenic lines of cv Cadenza expressing the 1Ax1 transgene when a whole plasmid or a clean fragment was used for transformation.

However, they did not study B73-6-1, which has a higher level of transgene expression and was found to show differences from the control line L88-6 in the present study.

The detailed analyses reported here demonstrate that it is possible to generate transgenic lines of wheat which are substantially equivalent to non-transformed lines in their grain metabolome. However, only a limited number of lines were analysed and one of these was found to differ when grown in four of six site–year combinations. Hence, the substantial equivalence of new lines needs to be estab-

lished on a case by case basis, and the metabolomic screening procedure described here provides a valuable tool for this process.

Experimental procedures

Materials

The production and characterization of the transgenic lines used in this study have been described previously (Barro *et al.*, 1997; Rooke *et al.*, 2003). T₃ seeds were planted in the field in 1998 and T₄, T₅ and T₆ seeds from the previous years' trials in 1999, 2000 and 2001. The T₅, T₆ and T₇ generations (harvested in 1999, 2000 and 2001) were used in the present study.

Field trials

Lines were planted in field experiments on two sites, LARS (near Bristol, UK) and RRes (Harpenden, UK), using a replicated randomized block design. Each experiment comprised

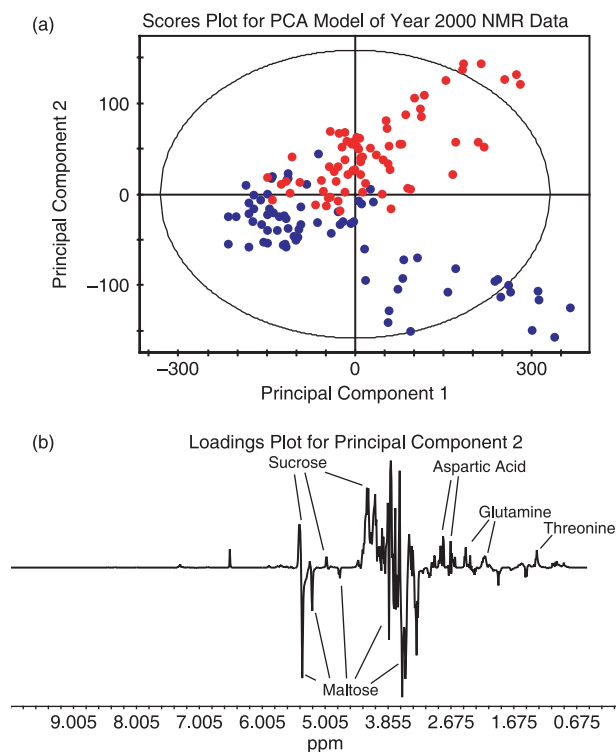


Figure 7 Use of loadings plots to identify metabolites contributing to differences between sites. (a) Scores plot from principal components analysis (PCA) model of year 2000 data only (blue ●, LARS; red ● circle, RRes). (b) Loadings plot from the same model showing the contribution of amino acids and carbohydrates to the site differences. LARS, Long Ashton Research Station; RRes, Rothamsted Research.

four blocks of 21 (seven \times three) 1-m² plots surrounded by a 2-m guard strip of a wheat cultivar which flowered at a different time from the trial wheat lines. Each plot contained about 350 plants and was separated by a 0.5-m path. At each site, the whole experiment was netted during establishment and grain filling, and was surrounded by either 20 m of a break crop or fallow land. The lines were randomly allocated to plots within each block, with a minimum of one plot of each line per block. Seeds were sown, harvested and threshed by hand. Consent to release these transgenic lines (licence number 97/R8/3) was obtained from the UK Department of the Environment Advisory Committee on Releases to the Environment under Part VI of the Environmental Protection Act 1990 and the Genetically Modified Organisms (Deliberate Release) Regulations 1992 as Amended (1995 and 1997).

Grain processing

One randomly selected plot of each line was selected from each block, and the grain from this plot was milled using a Brabender Quadrumat Junior (Duisburg, Germany) laboratory

mill to give white flour. Thus, in most cases, there were four flour samples for each of the six lines grown at the two sites for 3 years. However, in the 1999 field, only two samples of line L88-6 were available for each site. Similarly, only three samples of B102-1-1 were available from the 1999 Long Ashton trial.

¹H NMR sample preparation

Samples were prepared using a modified form of the protocol described by Ward *et al.* (2003). Replicate aliquots of white flour (30 mg) were weighed into 1.5-mL Eppendorf tubes. D₂O-CD₃OD (1 mL, 80 : 20) containing 0.05% w/v TSP-d₄ (sodium salt of trimethylsilylpropionic acid) was added to each sample. The contents of the tube were mixed thoroughly and heated at 50 °C in a water bath for 10 min. The samples were then spun down in a microcentrifuge for 5 min; 800 μ L of the supernatant was transferred to a clean Eppendorf tube and heated at 90 °C in a water bath for 2 min. The high-temperature (90 °C) step was incorporated to ensure that enzyme activity had stopped. The samples were then cooled at 4 °C for 45 min prior to re-centrifugation for 5 min (still at 4 °C); 700 μ L of the supernatant was transferred to a 5-mm NMR tube.

¹H NMR data collection

¹H NMR spectra were acquired under automation at a temperature of 300°K on a Bruker Avance (Coventry, UK) spectrometer operating at 399.752 MHz using a multinuclear, broad-band, 5-mm probe. The residual HOD signal was suppressed by pre-saturation and a relaxation delay of 5 s was employed. Each spectrum consisted of 1024 scans of 32K data points with a spectral width of 4845 Hz. The spectra were automatically Fourier transformed after the application of an exponential window function with a line broadening of 0.5 Hz. Phasing and baseline correction were carried out within the automation program. ¹H NMR chemical shifts were referenced to TSP-d₄ at δ 0.00.

Processing of ¹H NMR spectra

The ¹H NMR spectra were automatically reduced to ASCII files using AMIX (v3.0, Bruker Biospin, Coventry, UK). Spectral intensities were scaled to the TSP-d₄ peak and were reduced to integrated regions or 'buckets' of equal width (0.01 p.p.m.) corresponding to the regions between δ 9.000 and δ 0.500. The residual HOD and CD₂HOD peaks from δ 4.875 to δ 4.735 and from δ 3.355 to δ 3.185, respectively, were then excluded from all data sets. Unsupervised multivariate analyses by PCA were performed using SIMCA-P 9.0 (Umetrics, Umea, Sweden) employing mean centred scaling.

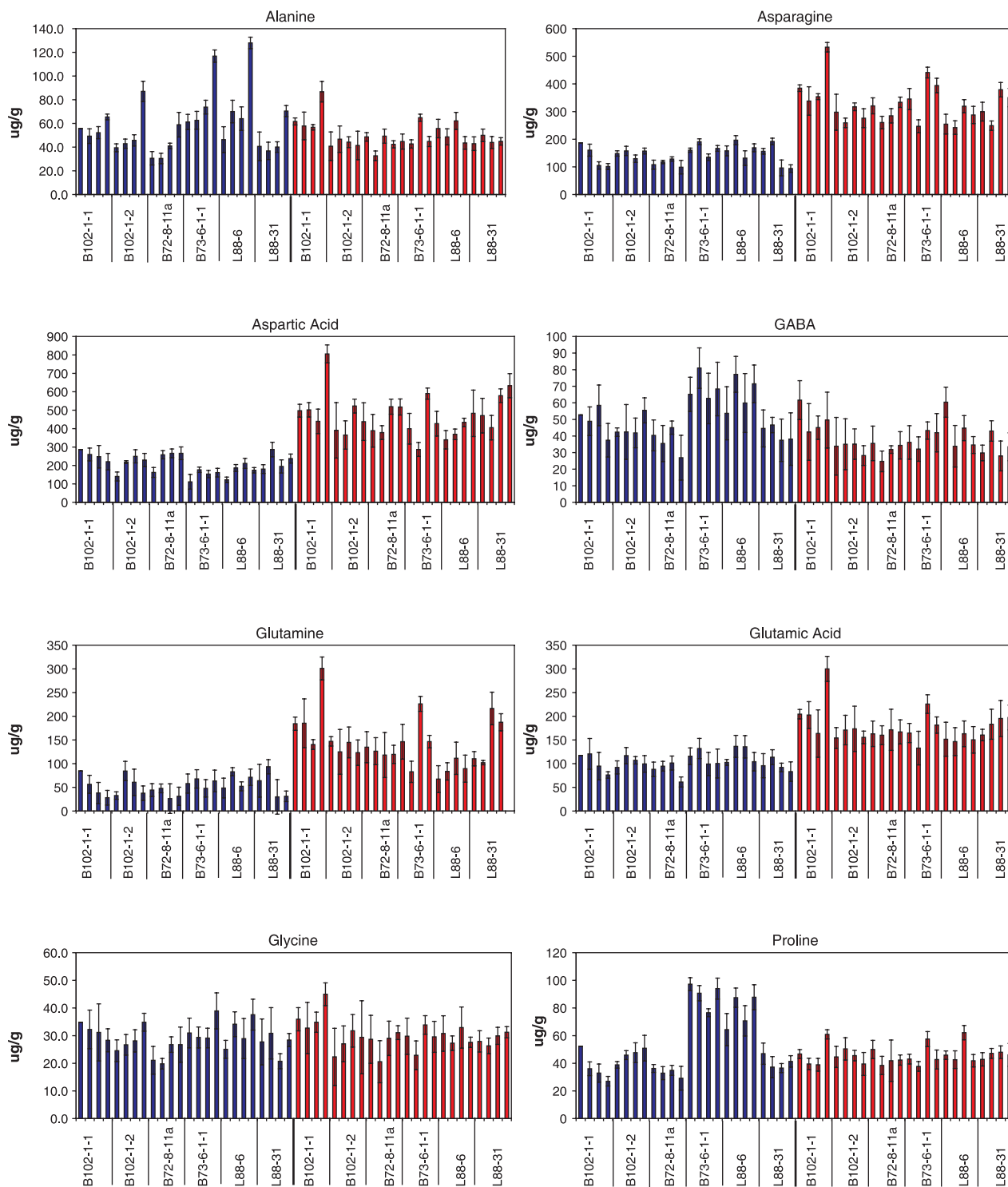


Figure 8 Concentrations of selected free amino acids in the plants from 48 plots grown in 2000, determined by gas chromatography-mass spectrometry (GC-MS). Red bars, Rothamsted Research (RRs); blue bars, Long Ashton Research Station (LARS). GABA, γ -aminobutyric acid.

Quantitative amino acid analysis by GC-MS

Five replicates from each flour sample from the 2000 field trials (48 samples arising from four plots of six lines at two sites) were analysed. The whole analysis was carried

out using a randomized block design consisting of five blocks of 4 days. Each day, 12 analytical samples and three calibration mixtures (or blank) were processed. Thus, one replicate analysis of each flour sample was performed every 4 days.

White flour (50 mg) was weighed into 2-mL Eppendorf tubes. Internal standard (1 mL of 100 µg/mL norvaline in water–methanol 80 : 20) and two tungsten carbide beads were added to each sample. The sample was agitated for 5 min at 30 Hz (Qiagen M300 mixer, Crawley, W. Sussex, UK). The samples were then spun down in a microcentrifuge for 5 min. The supernatant was transferred to a clean glass volumetric flask (5 mL). The extraction was repeated twice more, each with 1 mL of water–methanol (80 : 20), and the combined extract (c. 3 mL) was used for amino acid analysis.

Amino acid analysis was performed using the 'EZfaast' kit (Phenomenex, Macclesfield, Cheshire, UK) for sample preparation and derivatization, and GC-MS to identify and quantify the analytes. The kit is based around miniature cation exchange cartridges (built into Eppendorf tips) which were used to extract basic compounds from 200 µL of the extract. Once eluted from the cartridges, the basic compounds were derivatized with the propyl chloroformate reagent supplied with the kit prior to analysis by GC-MS (as described in the EZfaast manual).

GC-MS was performed on a Hewlett Packard 5890 Gas Chromatograph with a Hewlett Packard 5970 Series Mass Selective Detector (MSD) (now Agilent, S. Queensferry, West Lothian, UK) and a 10-m Zebron amino acid column (supplied with the EZFaast amino acid analysis kit). A column head pressure of 90 kPa was used. Two microlitres of derivatized sample were injected under split-less injection conditions. Both the injector and transfer line were held at 280 °C. After being held at 75 °C for 2 min, the oven temperature was ramped to 320 °C at 25 °C/min, with a further hold at this temperature for 1.2 min. The MSD scanned from 45 to 450 Da with a scan rate of 1.98 scans/s; the electron multiplier was set to 2600 V and the threshold was 150.

Quantification of the peaks in the chromatograms was achieved using extracted ion traces as opposed to the total ion chromatogram. Peak areas were calculated relative to the area of the norvaline extracted ion peak (*m/z* 158). The extracted ions used were as follows: alanine, 130; β-alanine, 129; asparagine, 155; aspartic acid, 216; γ-aminobutyric acid, 130; glutamine, 169; glutamic acid, 230; glycine, 116; isoleucine, 130; leucine, 172, 170; methionine, 203; ornithine, 156; phenylalanine, 190; proline, 156; serine, 146; threonine, 101; tryptophan, 130; valine, 158. Calibration curves for these amino acids were constructed to span the range of concentrations expected in the extracts. Calibration mixtures were run alongside each daily batch of samples.

Acknowledgements

Rothamsted Research receives grant-aided support from the

Biotechnology and Biological Sciences Research Council of the UK. This work was supported by the UK Food Standards Agency under the GO2 programme (grant number GO2003) 'Comparison of the metabolome and proteome of GM wheat: defining substantial equivalence'. Part of this work was supported by a grant to Long Ashton Research Station from Zeneca Ltd (Subsequently Syngenta) (1999–2002).

References

- Altpeter, F., Vasil, V., Srivastava, V. and Vasil, I.K. (1996) Integration and expression of high-molecular-weight glutenin subunit 1Ax1 gene into wheat. *Nat. Biotechnol.* **14**, 1155–1159.
- Altpeter, F., Baisakh, N., Beachy, R., Bock, R., Capell, T., Christou, P., Daniell, H., Datta, K., Datta, S., Dix, P.J., Fauquet, C., Huang, N., Kohli, A., Mooibroek, H., Nicholson, L., Nguyen, T.T., Nugent, G., Raemakers, K., Romano, A., Somers, D.A., Stoger, E., Taylor, N. and Visser, R. (2005) Particle bombardment and the genetic enhancement of crops: myths and realities. *Mol. Breed.* **15**, 305–327.
- Alvarez, M.L., Guelman, S., Halford, N.G., Lustig, S., Reggiardo, M.I., Ryabushkina, N., Shewry, P.R., Stein, J. and Vallejos, R.H. (2000) Silencing of HMW glutenins in transgenic wheat expressing extra HMW subunits. *Theor. Appl. Genet.* **100**, 319–327.
- Barcelo, P., Rasco-Gaunt, S., Thorpe, C. and Lazzeri, P.A. (2001) Transformation and gene expression. In *Advances in Botanical Research Incorporating Advances in Plant Pathology*, Vol. 34 (Shewry, P.R., Lazzeri, P.A. and Edwards, K.J., eds), pp. 59–126. London: Academic Press.
- Barro, F., Rooke, L., Bekes, F., Gras, P., Tatham, A.S., Fido, R.J., Lazzeri, P., Shewry, P.R. and Barcelo, P. (1997) Transformation of wheat with HMW subunit genes results in improved functional properties. *Nat. Biotechnol.* **15**, 1295–1299.
- Baudo, M.M., Lyons, R., Powers, S., Pastori, G.M., Edwards, K.J., Holdsworth, M.J. and Shewry, P.R. (2006) Transgenesis has less impact on the transcriptome of wheat grain than conventional breeding. *Plant Biotechnol. J.* **4**, 369–380.
- Blechl, A.E. and Anderson, O.D. (1996) Expression of a novel high-molecular weight glutenin subunit gene in transgenic wheat. *Nat. Biotechnol.* **14**, 875–879.
- Charlton, A., Allnut, T., Holmes, S., Chisholm, J., Bean, S., Ellis, N., Mullineaux, P. and Oehlschlager, P. (2004) NMR profiling of transgenic peas. *Plant Biotechnol. J.* **2**, 27–35.
- Choi, H.-K., Choi, Y.H., Verberne, M., Lefeber, A.W.M., Erkelens, C. and Verpoorte, R. (2004) Metabolic fingerprinting of wild type and transgenic tobacco plants by ¹H NMR and multivariate analysis technique. *Phytochemistry*, **65**, 857–864.
- Christensen, A.H. and Quail, P.H. (1996) Ubiquitin promoter-based vectors for high-level expression of selectable and/or screenable marker genes in monocotyledonous plants. *Trans. Res.* **5**, 213–218.
- Defernez, M., Gunning, Y.M., Parr, A.J., Shepherd, L.V.T., Davies, H.V. and Colquhoun, I.J. (2004) NMR and HPLC-UV profiling of potatoes with genetic modifications to metabolic pathways. *J. Agric. Food Chem.* **52**, 6075–6085.
- FAO/WHO (2000) *Safety Aspects of Genetically Modified Foods of Plant Origin. Report of a Joint FAO/WHO Expert Consultation on Foods Derived from Biotechnology*. Geneva, Rome: FAO.

- Garratt, L.C., Linforth, R., Taylor, A.J., Lowe, K.C., Power, J.B. and Davey, M.R. (2005) Metabolite fingerprinting in transgenic lettuce. *Plant Biotechnol. J.* **3**, 165–174.
- Jones, H.D. (2005) Wheat transformation: current technology and applications to grain development and composition. *J. Cereal Sci.* **41**, 137–147.
- Kohli, A., Twyman, R.M., Abranches, R., Wegel, E., Stoger, E. and Christou, P. (2003) Transgene integration, organization and interaction in plants. *Plant Mol. Biol.* **52**, 247–258.
- Kok, E.J. and Kuiper, H.A. (2003) Comparative safety assessment for biotech crops. *Trends Biotechnol.* **21**, 439–444.
- Kuiper, H.A., Kleter, G.A., Noteborn, J.M. and Kok, E.J. (2001) Assessment of the food safety issues related to genetically modified foods. *Plant J.* **27**, 503–528.
- Lawrence, G.J., MacRitchie, F. and Wrigley, C.W. (1988) Dough and baking quality of wheat lines deficient in glutenin subunits controlled by the *Glu-A1*, *Glu-B1* and *Glu-D1* loci. *J. Cereal Sci.* **7**, 109–112.
- Le Gall, G., Colquhoun, I.J., Davis, A.L., Collins, G.J. and Verhoeven, M.E. (2003) Metabolite profiling of tomato (*Lycopersicon esculentum*) using ¹H NMR spectroscopy as a tool to detect potential unintended effects following a genetic modification. *J. Agric. Food Chem.* **51**, 2447–2456.
- Manetti, C., Bianchetti, C., Bizzarri, M., Casciani, L., Castro, C., D'Ascenzo, G., Delfini, M., Di Cocco, M.E., Lagana, A., Miccheli, A., Motto, M. and Conti, F. (2004) NMR-based metabolomic study of transgenic maize. *Phytochemistry*, **65**, 3187–3198.
- McCann, M.C., Liu, K., Trujillo, W.A. and Dobert, R.C. (2005) Glyphosate-tolerant soybeans remain compositionally equivalent to conventional soybeans (*Glycine max* L.) during three years of field testing. *J. Agric. Food Chem.* **53**, 5331–5335.
- Oberdoerfer, R.B., Shillito, R.D., De Beuckeleer, M. and Mitten, D.H. (2005) Rice (*Oryza sativa* L.) containing the *bar* gene is compositionally equivalent to the nontransgenic counterpart. *J. Agric. Food Chem.* **53**, 1457–1465.
- OECD (1993) *Safety Evaluation of Foods Derived Through Modern Biotechnology: Concepts and Principles*. Paris: OECD.
- Payne, P.I. (1987) Genetics of wheat storage proteins and the effect of allelic variation on bread-making quality. *Annu. Rev. Plant Physiol.* **38**, 141–153.
- Rooke, L., Bekes, F., Fido, R., Barro, F., Gras, P., Tatham, A.S., Barcelo, P., Lazzeri, P. and Shewry, P.R. (1999) Overexpression of a gluten protein in transgenic wheat results in greatly increased dough strength. *J. Cereal Sci.* **30**, 115–120.
- Rooke, L., Steele, S.H., Barcelo, P., Shewry, P.R. and Lazzeri, P.A. (2003) Transgene inheritance, segregation and expression in bread wheat. *Euphytica*, **129**, 301–309.
- Sahrawat, A.K., Becker, D., Lutticke, S. and Lorz, H. (2003) Genetic improvement of wheat via alien gene transfer, an assessment. *Plant Sci.* **165**, 1147–1168.
- Shewry, P.R., Halford, N.G., Tatham, A.S., Popineau, Y., Lafandram, D. and Belton, P.S. (2003) The high molecular weight subunits of wheat glutenin and their role in determining wheat processing properties. *Adv. Food Nutr. Res.* **45**, 221–302.
- Shewry, P.R., Powers, S., Field, J.M., Fido, R.J., Jones, H.D., Arnold, G.M., West, J., Lazzeri, P.A., Barcelo, P., Barro, F., Tatham, A.S., Bekes, F., Butow, B. and Darlington, H. (2006) Comparative field performance over three years and two sites of transgenic wheat lines expressing HMW subunit transgenes. *Theor. Appl. Genet.* in press. doi: 10.1007/s00122-006-0279-1
- Ward, J.L., Harris, C., Lewis, J. and Beale, M.H. (2003) Assessment of ¹H-NMR spectroscopy and multivariate analysis as a technique for metabolite fingerprinting of *Arabidopsis thaliana*. *Phytochemistry*, **62**, 949–957.