Application of metabolite profiling to the identification of traits in a population of tomato introgression lines

S. A. Overy¹, H. J. Walker¹, S. Malone¹, T. P. Howard¹, C. J. Baxter², L. J. Sweetlove², S. A. Hill² and W. P. Quick¹,*

¹ Department of Animal and Plant Sciences, University of Sheffield, Sheffield S10 2TN, UK
² Department of Plant Sciences, University of Oxford, South Parks Road, Oxford OX1 3RB, UK

Received 13 July 2004; Accepted 1 November 2004

Abstract

Naturally occurring variation in wild species can be used to increase the genetic diversity of cultivated crops and improve agronomic value. Populations of introgression lines carrying wild species alleles afford an opportunity to identify traits associated with the introgressed regions, and facilitate characterization of the biochemistry and genetics underlying these phenotypes. Understanding plant metabolic pathways and the interactions between genes, phenotype, and environment is fundamental to functional genomics. Successful analysis of the complex network of plant metabolism requires analytical methods able to record information on as many metabolites as possible. Metabolite profiling is used to provide a snapshot of the metabolome in samples which differ in a known factor such as genetic background. Differences between the metabolite profiles can identify those metabolites/metallic pathways affected by the introgression and allow genetic maps for metabolic alterations to be established. A Time-of-Flight Mass Spectrometry method is presented, with associated data reduction, used for profiling aqueous metabolites from tomato. Analysis of ripe fruits of two tomato species, Lycopersicon esculentum and L. pennellii, showed differences in the amounts of many metabolites, including organic acids and sugars. Six introgression lines, L. pennellii introgressions within L. esculentum, were also examined and showed that Principal Component Analysis can reveal subtle differences in metabolism of the introgressed lines when compared to their parents.

Key words: Introgression, profiling, tomato.

Introduction

Attempts to manipulate plant metabolism by genetic modification do not always yield successful, or predicted, outcomes. Endogenous regulation of metabolic flux can result in the effects of the transgene being absorbed, producing an essentially unchanged, or even undesirable, phenotype (Morandini and Salamini, 2003). In addition, the genetic basis underlying many desirable traits, such as yield, is complex, further complicating attempts at manipulation. Therefore, there is a clear requirement for better understanding of the regulation of plant metabolism, the interactions between genes and the environment, and the relationship between genes and phenotype. As an alternative to the candidate gene approach, novel traits may be introduced into plants via chromosomal introgressions from related species. Subsequent analyses of the resulting phenotypes can further inform the genetics underlying these traits, and may allow the discovery of new metabolic pathways.

Centuries of plant breeding and domestication have resulted in modern crops with a very narrow genetic base. Chromosomal introgressions from related varieties or wild species can be used to increase the genetic diversity of the crop without losing the desirable elite phenotype (Zamir, 2001). Introgression lines (ILs) are produced by crossing an elite crop variety with a species containing desirable germplasm, followed by successive backcrossing to the elite parent. Zamir and colleagues have produced a population of introgression lines of the cultivated tomato Lycopersicon esculentum (M82) containing defined chromosomal introgressions from the wild tomato species L. pennellii (Eshed and Zamir, 1995). Introgressions representing the entire tomato genome are covered in 50 lines (with an additional 26 lines containing sub-regions), with each line homozygous for a single introgressed region.
Studies using tomato ILs have identified a number of quantitative trait loci (QTL) for improved agronomic traits such as fruit soluble solids, fruit weight and yield (Eshed and Zamir, 1995), and fruit colour (Liu et al., 2003), plus other traits such as undesirable volatiles (Tadmor et al., 2002), and leaf dissection (Holtan and Hake, 2003). In some cases, the genetic basis for the QTL has been mapped to a single gene, and variation is displayed in both coding and regulatory regions (Paran and Zamir, 2003). For example, a QTL on chromosome 9, Brix 9-2-5, results in fruits with increased brix (soluble solids) and no reduction in yield. The QTL has been delimited to a 484 bp region of an apoplastic invertase gene Lin5, expressed in tomato fruits and flowers (Fridman et al., 2000; Fridman and Zamir, 2003). fw2.2, a QTL controlling fruit weight, has been mapped to a small (<150 kb) region of chromosome 2 (Alpert and Tanksley, 1996), while the B locus on chromosome 6 has been identified as a novel lycopene β-cyclase producing an increase in β-carotene in ripe fruits (Ronen et al., 2000). QTL for different traits have been shown to co-localize on the same area of the chromosome, for example, loci controlling fruit shape, weight, and epidermal reticulation on chromosome 4 (Yates et al., 2004), which could indicate multiple effects of a single gene mutation.

Clearly the breeding approach has benefits in transferring desirable traits to domesticated species. The use of a stable population of ILs means that a wide range of traits can be studied under a variety of environmental conditions, thereby amassing a large amount of information on phenotypes and gene/environment interactions. As outlined above, several studies have demonstrated the transfer of beneficial agronomic traits to the domesticated variety. There is, however, little information on how the wild species introgressions have affected the biochemical phenotype of the crop species. Detailed examination of the biochemical and molecular basis underlying these traits will advance the understanding of metabolic pathways, their interactions with the environment, and their contribution to quantifiable traits.

Plant metabolic pathways are increasingly viewed as a network rather than a series of linear reactions, and it is therefore important when studying perturbations in a particular pathway to look beyond the immediately connected metabolites (Sweetlove et al., 2003; Trethewey, 2004). With this requirement comes the need for robust analytical methods, able to detect a large number of chemically diverse metabolites. Analytical approaches to the analysis of the metabolome vary from targeted analyses, concentrating primarily on a single class of metabolites such as amino acids or sugars, to more non-targeted approaches where the aim is to provide information on as comprehensive as possible a selection of metabolites. Functional genomic strategies would ideally involve the measurement of all metabolites within a sample in order to investigate pleiotropic effects. Currently, however, no analytical technique can provide detection of all metabolites in one sample, and so truly non-biased analyses are not available. Directed metabolite analyses require the user to determine in advance which metabolites are of interest and are often time intensive. Metabolite fingerprinting, or footprinting, can be used to depict a ‘semi-quantitative’ snapshot of the metabolome at a given time. Profiles generated from samples separated by temporal, environmental or genetic factors can be compared. Metabolite fingerprinting techniques have been used, in combination with chemometrics, for a range of applications, including examination of responses of plants to environmental stress (Johnson et al., 2003) and plant–plant interference (Gidman et al., 2003), discrimination of vegetable and nut oils (Goodacre et al., 2002), classification of yeast mutants (Allen et al., 2003), and identification of bacterial strains (Vaidyanathan et al., 2002).

A method is presented for generating metabolite profiles using Electrospray Ionization Time-of-Flight Mass Spectrometry (ESI-TOF-MS), with associated data reduction. The high mass resolution of TOF instruments allows the detection of metabolites of different monoisotopic masses, but the same nominal mass, for example, glutamine 146.0691Da and lysine 146.1055Da.

Metabolite profiling is being used as part of these investigations into the phenotypes resulting from the chromosomal introgressions, with a view to mapping metabolic differences to the introgressed regions.

Materials and methods

Plant growth

Seeds of L. esculentum M82, L. pennellii and the introgression lines were obtained from the Tomato Genetic Resource Centre, University of California, Davis. Eight plants of each line (except IL3-2 where poor germination meant only three plants were included in this study) were grown in pots in a glasshouse under ambient conditions, in Levington M3 compost with supplemented nutrients (Osmocote, Scotts Ltd, UK). Plants were randomized throughout the glasshouse in order to minimize the effects of local environmental conditions. Plants were watered daily and fed weekly with a proprietary tomato food (Tomato liquid, LBS Horticulture, Lancashire, UK) after flowering.

For measurement of leaf area, three plants of each line were harvested at the start of the flowering stage. Leaves were stripped and leaf area recorded using a (Delta-T Devices Ltd, Cambridge, UK).

The remaining five plants (three for IL 3-2) were allowed to fruit. Ten fruit from each plant were used for measurement of brix. Brix was measured on the homogenized pericarp of ripe fruit using a portable refractometer (Bellingham and Stanley Ltd, Kent, UK). At the time of final harvest, individual fruits were weighed and the yield taken as the sum of the fresh weight of all the fruit from that plant (including non-ripe fruit). One plant of IL 5-4 did not produce any fruit and therefore only four plants were used for this line.

Metabolite sample extraction and preparation

Samples were taken from the pericarp of ripe fruits of L. esculentum and L. pennellii, and were immediately frozen in liquid nitrogen after harvesting. For all lines, ripeness was judged as being 7–10 d post-breaker (i.e. after the first appearance of carotenoid colouration). For
metabolite profiling, samples were extracted from three fruits per plant, three plants per line. Samples were extracted in methanol/ chloroform based on Valle et al. (1998). 0.1–0.2 g of pericarp was ground in 750 µl of an extraction medium comprising water/ chloroform/methanol (8/20/47, by vol.), and left on ice for 30 min. Aqueous metabolites were extracted twice with 400 µl water (polar extract) and stored at −80 °C until analysis.

**Direct-injection ESI-TOF-MS**

ESI-TOF-MS was performed with an LCT mass spectrometer (Waters Ltd, Manchester, UK). MassLynx data system (version 4.0) running under Windows NT on an IBM-compatible PC provided instrument control, data acquisition, and data processing. The mass spectrometer was operated in negative ion mode at a resolution of 4000 (FWHM) at mass 200 m/z and at a capillary voltage of 2800 V, extraction cone at 5 V, and sample cone at 20 V. An RF lens of 75 V was chosen to allow optimum detection of lower mass ions. Source and desolvation temperatures were maintained at 110 °C and 120 °C, respectively, with desolvation and nebulizer gases at flow rates of 400 and 100 l h⁻¹, respectively. Spectra were collected in centroid mode in the mass range 50–800 Da at a rate of one spectrum s⁻¹ (0.95 s scan time, 0.05 s interscan delay) and 180 spectra were summed over a period of 3 min before further data analysis. No background subtraction was performed. Mass calibration was performed daily in the range 50–800 m/z with organic acids and polyethylene glycol.

---

![L. esculentum parent](image1.png) ![IL 3-2](image2.png) ![IL 4-4](image3.png)

**Fig. 1.** Phenotypes observed for fruit of *L. esculentum* M82 and two of the introgression lines grown in the glasshouse. Scale 1 cm=2.2 cm.

![Graph A](graph_a.png) ![Graph B](graph_b.png) ![Graph C](graph_c.png) ![Graph D](graph_d.png)

**Fig. 2.** Some indicators of performance of glasshouse-grown plants in the experiment. (A) Leaf area, (B) average fruit brix, (C) yield, and (D) horticultural yield (product of brix and yield). Data are from the means of three plants per line for leaf area and five plants per line for fruit measurements (except IL 3-2 and IL 5-4 where data are from three and four plants, respectively), and expressed as a percentage of the mean value for the *L. esculentum* parent. Black bars indicate a significant difference from the *L. esculentum* parent (*P* <0.05).
diacid. The use of organic acids improved the accuracy of mass calibration in the mass range 60–200 m/z, where many metabolites were detected. A Lockspray\textsuperscript{TM} interface was used to allow automated accurate mass measurements of plant metabolites and for daily quality control checks. Leucine enkephalin, 5 \( \mu \)g cm\(^{-3}\) in 1/1 (v/v) acetonitrile/water, was employed during Lockspray operation as an internal mass reference. A syringe pump (Razel, Connecticut, USA) at a flow rate of 10 \( \mu \)l min\(^{-1}\) was used to introduce samples directly into the mass spectrometer. Samples were analysed in a randomized order to minimize effects of any day-to-day machine variation.

**Data analysis**

Peak lists (mass to four decimal places versus ion counts) were transferred from the MassLynx data system to Microsoft Excel (Microsoft Corp, USA) as text files. Many metabolites are present in relatively low abundance, so removal of background noise by imposition of a peak threshold could result in loss of peaks of interest. Instead, samples were run in triplicate and the spectra combined to select only those peaks present in all three replicates. Peaks are selected if the mass variance between the three replicates of a recorded mass falls within an accepted range. It was found that defining the acceptable mass variance as a linear function of the m/z generates the maximum number of peaks with minimum false positives. The following equations were used: for ES+, \( y < 0.00003x + 0.0033 \); for ES\textsuperscript{-}, \( y < 0.00003x + 0.0044 \), where \( y \) is the standard deviation of the three masses and \( x \) is the mean of the three masses. Once a peak is selected across the three analytical replicates, the mean mass is calculated and used as the mass for all three runs. These masses, with associated intensities, form the metabolite profile for that sample.

---

**Fig. 3.** Metabolite profiling of two tomato species. Representative spectra are shown for aqueous extracts, negative ionization, from ripe fruits of (A) *L. esculentum* and (B) *L. pennellii*.
To minimize variation between different extractions and analytical replicates, the data were normalized to the total ion count (TIC) in each replicate.

For ease of comparison, in this study data were rounded into 1 mass unit ‘bins’ and the abundances (as percentage TIC) summed for each bin. Samples were then compared by Principal Component Analysis (PCA) using SPSS (v.12.0, SPSS Inc., Illinois, USA). Differences were tested for significance using ANOVA (Minitab 13.31, Minitab Inc.).

Results and discussion

Glasshouse-grown plants exhibit many of the phenotypes seen in the field. Figure 1 shows phenotypes observed for two of the ILs. IL3-2 has the yellow fruit characteristic of the yellow-flesh \( (r) \) mutation, attributed to a null mutation in \( \text{Psy-1} \) (Fray and Grierson, 1993). IL4-4 exhibits the smaller fruit and epidermal reticulation seen in other studies (Monforte \textit{et al.}, 2001; Yates \textit{et al.}, 2004), and possibly indicates a pleiotropic phenotype. The visible phenotypes of the fruits of the other ILs discussed in this paper did not differ from that of the parent line and are therefore not included in Fig. 1.

Figure 2 shows some indicators of plant performance in the glasshouse for selected ILs. None of the lines presented showed a significant difference in leaf area (Fig. 2A), indicating that canopy size should not be a contributing factor to differences in fruit yield or brix. Brix is a measure of soluble solids, which in tomato fruit is mainly sugars and organic acids, both major contributors to fruit quality and flavour. Several of the ILs showed an increase in fruit brix (Fig. 2B), although this was only significant (\( P < 0.05 \)) in one line (IL 4-4). All the lines with increased brix also showed reduced yield (Fig. 2C), resulting from either smaller fruit (IL 4-4) or decreased number of fruit (IL 3-2, IL 7-3, and IL 5-4; data not shown). An interesting observation with one plant of IL 4-4 was that two fruit which did not show the characteristic epidermal reticulation, but were still smaller than parent fruit, also had a lower brix (around the level of that in the parent; data not shown), suggesting that the reticulation and brix phenotypes may be linked. IL 4-3 was the only line to show an increase in yield and a corresponding increase in horticultural yield (product of brix and yield; Fig. 2D), although again this increase was not statistically significant. Generally, attempts to breed tomato plants with higher fruit solids have not been successful, due to the negative relationship between yield and brix (Allen Stevens and Rick, 1986). In addition, environmental factors, such as irrigation and disease resistance, have a large impact on fruit solids (Allen Stevens and Rick, 1986). In field trials, differences in yield and horticultural yield were observed between plants grown in plots and those grown individually, attributed to competition between plants (Eshed \textit{et al.}, 1996). In addition, plot-grown plants (35 plants per 10 m\(^2\)) showed higher experimental variability, especially in yield and plant weight, meaning that only very high differences were statistically significant (Eshed and Zamir, 1994). The data presented above show some similarities to those found for field-grown plants, with significant increases in brix in ILs 4-4, 5-4, and 7-3, and a significant reduction in yield in IL 5-4 (Eshed and Zamir, 1995). However, there were also some differences in the two studies. In the field-grown crops, increases in brix were also seen in ILs1-4 and 3-2, while none of the above ILs showed a significant change in

<table>
<thead>
<tr>
<th>Compound</th>
<th>Molecular formula</th>
<th>Exact mass</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pyruvic acid</td>
<td>C(_3)H(_4)O(_3)</td>
<td>88.0160</td>
</tr>
<tr>
<td>Aminobutyric acid</td>
<td>C(_4)H(_6)N(_2)O(_2)</td>
<td>103.0633</td>
</tr>
<tr>
<td>Maleic acid/fumaric acid</td>
<td>C(_4)H(_4)O(_4)</td>
<td>116.0110</td>
</tr>
<tr>
<td>Succinic acid</td>
<td>C(_4)H(_6)O(_3)</td>
<td>118.0266</td>
</tr>
<tr>
<td>Homo-serine</td>
<td>C(_5)H(_9)NO(_2)</td>
<td>119.0580</td>
</tr>
<tr>
<td>Cysteine</td>
<td>C(_3)H(_7)NO(_2)S</td>
<td>121.0196</td>
</tr>
<tr>
<td>Aspartic acid</td>
<td>C(_4)H(_7)NO(_4)</td>
<td>133.0375</td>
</tr>
<tr>
<td>Malic acid</td>
<td>C(_4)H(_6)O(_4)</td>
<td>134.0215</td>
</tr>
<tr>
<td>Glutamic acid</td>
<td>C(_4)H(_8)NO(_4)</td>
<td>147.0532</td>
</tr>
<tr>
<td>Cinnamic acid</td>
<td>C(_5)H(_8)NO(_2)</td>
<td>148.0524</td>
</tr>
<tr>
<td>Methionine</td>
<td>C(_5)H(_10)NO(_2)S</td>
<td>149.0508</td>
</tr>
<tr>
<td>Pentose sugars(^a)</td>
<td>C(_5)H(_7)O(_2)</td>
<td>150.0525</td>
</tr>
<tr>
<td>Histidine</td>
<td>C(_6)H(_9)N(_3)O(_2)</td>
<td>155.0693</td>
</tr>
<tr>
<td>Coumaric acid/hydroxycinnamic acid/phenylpyruvate</td>
<td>C(_6)H(_8)O(_3)</td>
<td>164.0473</td>
</tr>
<tr>
<td>Hexose sugars</td>
<td>C(_6)H(_10)O(_5)</td>
<td>180.0630</td>
</tr>
<tr>
<td>Citric acid</td>
<td>C(_6)H(_8)O(_7)</td>
<td>192.0270</td>
</tr>
<tr>
<td>Sedoheptulose</td>
<td>C(_7)H(_10)O(_7)</td>
<td>210.0739</td>
</tr>
</tbody>
</table>

\(^a\) Or fragment ions from higher molecular weight sugars.
horticultural yield. These differences may reflect environmental modulation of the QTL.

It is encouraging that many of the traits observed in field-grown crops, under different environmental and temporal conditions, are replicated in the glasshouse-grown plants. Other studies in tomato have also shown consistency of selected QTL effects across different environments (Monforte et al., 2001). This relative stability in the expression of these QTL should hopefully mean that metabolic traits identified in the plants can also be reproducibly observed.

**Metabolite profiling: data processing and reduction**

For a negative ion analysis of a tomato fruit aqueous extract, a single ‘metabolite profile’ contains around 4600 mass peaks, potentially comprising metabolite ‘molecular’, fragment and isotope ions, plus matrix-associated peaks.

---

Fig. 5. Proportion of the total ion count found in selected bins in the two tomato species. Normalized peak areas were summed for each bin. Data are means of eight replicates for each species ± 1 standard deviation. Significant (P <0.05) differences are marked with an asterisk. (A) Bin 133, (B) Bin 179, (C) Bin 191, (D) Bin 215, (E) Bin 281, (F) Bin 487.
and machine noise. To eliminate the noise from the profile, a minimum threshold is often used, which assumes that noise peaks, due to their random nature, will be smaller than metabolite peaks. However, this is not necessarily the case, and it was found that the spectra are ‘noisier’ at the higher mass end of this study’s scan range, so eliminating these peaks on an intensity basis would discriminate against lower mass peaks. Therefore, in order to minimize the loss of low intensity metabolite peaks (which may reflect either naturally low abundance or the presence of a poorly ionized compound) an alternative approach to noise reduction has been adopted. As stated previously, each sample is analysed in triplicate and if a peak is determined to be present in all three analytical runs it is retained in the peak list. Peaks which cannot be detected within a defined mass range across all three replicates are discarded and not included in subsequent analysis. For peaks which are retained, the mean mass across the analytical replicates is calculated and used as the mass for that sample.

This data-reduction step results in a loss of around 50% of the peaks, with a corresponding reduction in the TIC to around 87% of the original (data not shown). Most of the discarded peaks are of low abundance, or greater than c. 600 m/z where the spectra are more ‘noisy’. In order to minimize sample-to-sample variation, mass intensities were normalized to the TIC before further analysis.

Metabolite profiling of two tomato species

Data are presented from negative ion analysis of aqueous extracts of fruits of *L. esculentum* and *L. pennellii*. Profiles for both species are data-rich below 350 m/z (corresponding to the mass range of most primary metabolites), and the most intense peaks can be identified on the basis of their mass as citric acid ([M-H]− 191.0192) and malic acid ([M-H]− 133.0137), plus other as yet unidentified masses (Fig. 3). In addition, the C13 isotope peaks for these compounds are prominent. However, it should be remembered that peak intensity is not necessarily a reflection of the relative concentrations of these compounds in the sample, since differences in ionization efficiency greatly influences the relative size of the peaks. Candidate identities, based on measured mass of [M-H]−, for some of the peaks found in the profiles are shown in Table 1.

For ease of comparison, the data have been summed into 1 mass unit ‘bins’. This has obvious disadvantages, notably the loss of mass identity and the inclusion of more than one metabolite within a bin, leading to the potential for large peaks to mask changes in smaller peaks in the same bin. However, it is also a relatively simple way of comparing the profiles of samples where large numbers of masses are observed, and where the masses are not exactly identical in each run.

The binned data from the two species were analysed by PCA using SPSS. Figure 4 shows clear discrimination of the two species in PC1. Many bins contribute to this separation (i.e. had high PCA loadings), a few of which are shown in Fig. 5. The bins shown in Fig. 5 have been selected to show that differences could be detected in bins covering a wide range of percentage TIC. In general, each 1 mass unit bin will contain more than one mass peak, which, as mentioned previously, may complicate identification of actual metabolites which change. Further examination of bins 133, 191, and 179 suggests that malic acid, citric acid, and hexose sugars ([M-H]− 179.0552) respectively are

![Fig. 6.](image-url)
responsible for the differences in those bins, with *L. pennisetii* having a greater proportion of TIC in the acids and less in hexose sugars.

**Metabolite profiling of tomato introgression lines**

Since it was possible to distinguish between the metabolite profiles of the two parent species, selected ILs were also examined. Figure 6A shows a PCA plot for six introgression lines and the *L. esculentum* parent. It can be seen that, in general, the biological variability was high, but there was, nevertheless, some discrimination of one of the ILs (IL 3-2) from the *L. esculentum* parent (shown more clearly in Fig. 6B). Examination of some of the metabolite bins contributing to PC1 shows that differences could be observed in the size of several of these bins, even though variation was high, and, as with the parent lines, differences were detected over a wide range of percentage TIC values (Fig. 7).

As stated previously, few studies to date have examined the influence of wild species introgressions on fruit metabolism. Fulton *et al.* (2002) examined four populations of

---

Fig. 7. Proportion of TIC found in selected bins in the introgression lines. Normalized peak areas were summed for each bin. Data are means of 9 replicates for each line (except M82 eight replicates) ± 1 standard deviation. Black bars indicate a significant (*P* <0.05) difference from *L. esculentum* M82 parent. (A) Bin 111, (B) Bin 150, (C) Bin 165, (D) Bin 191, (E) Bin 239, (F) Bin 787.
tomato plants, each carrying introgressions from a different wild species (L. hirsutum, L. peruvianum, L. parviflorum, and L. pimpinellifolium) for QTL affecting sugars and organic acids. They identified several QTL for the content of metabolites that contribute to fruit flavour—citric acid, malic acid, glutamic acid, glucose, fructose, and sucrose. Wild parent alleles were associated with an increase in acidity-related traits in 51% of the QTL, and concentrations of the acids were frequently correlated indicating either pleiotropic effects or the presence of clusters of QTL affecting multiple acids. One of these QTL for increased citric acid content, ca3.1/PM, is contained on the introgressed region within IL 3-2, and this line does appear to show an increase in normalized peak intensity for citric acid (Bin 191, Fig. 7D).

Despite the speed and relative non-selectivity of the direct injection method, a drawback is the inability to resolve structural isomers, such as the hexose sugars, a major component of tomato fruit. Direct infusion also raises concerns with ionization suppression and limited dynamic range, and the process of binning data results in loss of mass resolution and may lead to problems with larger peaks masking changes in smaller ones within the same bin. Chromatographic separation of the extracts prior to MS detection would resolve many of the above issues; however, many current LC protocols are lengthy and dependent on specific chemical groups of metabolites. LC-MS methods are currently being developed in order to resolve structural isomers better, whilst maintaining a relatively high sample throughput.

Concluding remarks

Direct-infusion ESI-TOF-MS are being used to generate metabolite profiles from tomato fruit extracts. Removal of noise peaks by selection of only those masses present in three analytical replicates has allowed the number of peaks, and therefore data complexity, in the profiles to be significantly reduced. Mass peaks can be assigned candidate identification due to the high mass resolution obtained with time-of-flight analysis.

Analysis of the metabolite profiles by PCA results in clear discrimination of the two tomato species. Further examination of the peaks underlying this clustering reveals differences in the levels of many metabolites, including citric acid, malic acid, and hexose sugars. Studies with six of the introgression lines suggest that any changes in metabolite profiles of the ripe fruits are relatively subtle, although at least one of the introgression lines does cluster apart from the L. esculentum parent in PCA. It may be that use of a supervised clustering method, such as discriminant function analysis (Goodacre et al., 2003; Johnson et al., 2003; Allen et al., 2003), would be better able to distinguish between the introgression lines and identify key differences. In addition, only a relatively small dataset has been presented here; data were also collected for positive ion and non-polar phases. Further work is ongoing to profile metabolites from the entire IL population, with a view to mapping any phenotypic differences to the introgressed regions. These data can be used in combination with other studies to advance understanding of plant systems biology, essential for successful functional genomics (Gur et al., 2004).

Acknowledgement

The authors would like to acknowledge financial support from the BBSRC.

References


Alpert KB, Tanksley SD. 1996. High-resolution mapping and isolation of a yeast artificial chromosome contig containing fw2.2: a major fruit weight quantitative trait locus in tomato. Proceedings of the National Academy of Sciences, USA 93, 15503–15507.


