Journal of
Experimental
Botany
www.jxb.oxfordjournals.org

RESEARCH PAPER

Metabolic profiling of strawberry (*Fragaria×ananassa* Duch.) during fruit development and maturation

Juanjuan Zhang[1], Xin Wang[1], Oliver Yu[2,3], Juanjuan Tang[1], Xungang Gu[2], Xiaochun Wan[2] and Congbing Fang[1,2,*]

[1] School of Horticulture, Anhui Agricultural University, Hefei 230036, Anhui, PR China
[2] Key Laboratory of Tea & Medicinal Plants and Safe Production, Ministry of Agriculture, Anhui Agricultural University, Hefei 230036, Anhui, PR China
[3] Donald Danforth Plant Science Center, St Louis, MO 63132, USA

* To whom correspondence should be addressed.E-mail: fcb_ah@ahau.edu.cn

Abstract

Strawberry (*Fragaria×ananassa* Duch), a fruit of economic and nutritional importance, is also a model species for fleshy fruits and genomics in Rosaceae. Strawberry fruit quality at different harvest stages is a function of the fruit's metabolite content, which results from physiological changes during fruit growth and ripening. In order to investigate strawberry fruit development, untargeted (GC-MS) and targeted (HPLC) metabolic profiling analyses were conducted. Principal component analysis (PCA) and partial least squares discriminant analysis (PLS-DA) were employed to explore the non-polar and polar metabolite profiles from fruit samples at seven developmental stages. Different cluster patterns and a broad range of metabolites that exerted influence on cluster formation of metabolite profiles were observed. Significant changes in metabolite levels were found in both fruits turning red and fruits over-ripening in comparison with red-ripening fruits. The levels of free amino acids decreased gradually before the red-ripening stage, but increased significantly in the over-ripening stage. Metabolite correlation and network analysis revealed the interdependencies of individual metabolites and metabolic pathways. Activities of several metabolic pathways, including ester biosynthesis, the tricarboxylic acid cycle, the shikimate pathway, and amino acid metabolism, shifted during fruit growth and ripening. These results not only confirmed published metabolic data but also revealed new insights into strawberry fruit composition and metabolite changes, thus demonstrating the value of metabolomics as a functional genomics tool in characterizing the mechanism of fruit quality formation, a key developmental stage in most economically important fruit crops.

Key words: *Fragaria×ananassa*, fruit maturation, fruit quality, GC-MS profiling, metabolic profiling, metabolite composition, metabolomics, strawberry.

Introduction

Strawberry (*Fragaria×ananassa* Duch.) is one of the most economically important fresh and processed fruits, consumed for both its pleasant flavour and its nutrient content (Hancock, 1999). Strawberry is cultivated throughout the world; ~4.07 million tons were produced in 2008 (http://faostat.fao.org/site/567/default.aspx).

Strawberry flavour is a result of a complex mixture of numerous volatile and organoleptic compounds combined with characteristics such as texture and taste. More than 300 volatile compounds have been identified in ripening strawberry (Honkanen and Hivi, 1990; Latrasse, 1991), which can be grouped into several chemical classes, including the major components acids, aldehydes, ketones, alcohols, esters, and lactones, and the contributing groups of sulphur compounds, acetals, furans, phenols, epoxides, and alkanes (Zabetakis and Holden, 1997). Although the individual volatiles of these groups are often present in minute quantities, typically in concentrations of 10–100 ppm

of the fruit fresh weight (Maarse, 1991), they have a significant impact on the overall flavour of strawberry fruits.

The nutritional quality of strawberry fruits is closely correlated with the presence of soluble sugars, organic acids, amino acids, and some major secondary metabolites. These compounds play an important role in maintaining fruit quality and nutritive value; for this reason, fruit compositional analysis is of interest to food chemists and processors. Phenolic acids and their derivatives are often conjugated with sugars and have been frequently reported in strawberry fruits and leaves (Määttä-Riihinen *et al.*, 2004; Aaby *et al.*, 2007; Hukkanen *et al.*, 2007). Anthocyanins are a major class of polyphenols producing fruit pigmentation and are strong antioxidants that contribute to the health-beneficial effect of strawberry. (Santos-Buelga and Scalbert, 2000; Gu *et al.*, 2003). Strawberry fruits possess high *in vitro* antioxidative activity, which has been positively correlated with the content of polyphenolic compounds and, specifically, anthocyanins (Heinonen *et al.*, 1998; Wang and Jiao, 2000; Wang and Lin, 2000).

The physiological changes that contribute to strawberry fruit quality during development and maturation result from changes in gene expression and enzyme activities (Medina *et al.*, 1997; Manning, 1998; Nam *et al.*, 1999). The most noticeable changes involve alterations to fruit shape, size, texture, and pigmentation that coincide with an increase in the content of soluble solids and the production of natural aroma and flavour compounds (Perkins-Veazie, 1995). So far, great efforts have been made to investigate fruit composition and to explain the relationship between fruit quality and numerous factors, including, but not limited to, strawberry varieties (Wang and Lin, 2000; Määttä-Riihinen *et al.*, 2004; Atkinson *et al.*, 2006), developmental stage (Wang and Lin, 2000;), degree of maturity (Perez *et al.*, 1992; Perkins-Veazie, 1995), climatic conditions (Anttonen *et al.*, 2006), cultivation practices (Zabetakis and Holden, 1997; Anttonen *et al.*, 2006; Atkinson *et al.*, 2006), and biotic or abiotic stresses (Terry *et al.*, 2007). These studies focused on either individual or just a few physiochemical properties, but failed to characterize the global changes to fruit biochemical composition.

Metabolomic approaches have increasingly been used to gain insight into the metabolic composition of plant organs and to characterize the natural variance in metabolite content (Schauer and Fernie, 2006). Metabolomics can provide a diagnostic tool for better understanding of a biological system and has now been successfully performed on a diverse array of plant species, including models such as *Arabidopsis* (Kim *et al.*, 2007), *Medicago* (Chen *et al.*, 2003; Broeckling *et al.*, 2005), and tobacco (Blount *et al.*, 2002), and the important crops tomato (Roessner-Tunali *et al.*, 2003; Schauer *et al.*, 2005), potato (Roessner *et al.*, 2001), rice (Sato *et al.*, 2004), wheat (Hamzehzarghani *et al.*, 2005), cucumber (Tagashira *et al.*, 2005), and strawberry (Aharoni *et al.*, 2002; Fait *et al.*, 2008). In contrast to transcriptomics and proteomics, which rely to a great extent on genome information, metabolomics is mainly metabolite dependent. The cultivated varieties of commercial strawberries, usually designated as *F.×ananassa*, are almost all octoploids (Hancock, 1999); therefore, the whole genome sequence of strawberry is still unavailable, which hinders genetic research based on transcriptomics and proteomics (Aharoni *et al.*, 2000). Consequently, it is essential to apply metabolomic approaches in strawberry research, which seems to have at least three advantages. First, strawberry fruit quality at harvest is a direct function of metabolite content, and the compositional analysis of plant metabolites is an established application in metabolomic research (Moco *et al.*, 2007). Secondly, metabolomics relies on the analysis of the multitude of small molecules (metabolites) present in a biological system (Schauer and Fernie, 2006; Moco *et al.*, 2007), which is helpful for the comprehensive understanding of the correlation between all the metabolites investigated and their metabolic network, instead of just a few. Thirdly, without the presence of genome information, metabolomics is a key tool in comprehensive functional genomics that aims to decipher gene function, investigate metabolic regulation, and analyse the systemic response to environmental or genetic perturbations (Schauer and Fernie, 2006).

Aharoni *et al.* (2002) analysed four consecutive stages (green, white, turning, and red) of strawberry fruit development and identified the changes of fruit metabolites by use of Fourier transform ion cyclotron mass spectrometry. Metabolic networks were reconfigured between the achene and the receptacle to investigate the cross-talk between primary and secondary metabolism during strawberry fruit development (Fait *et al.*, 2008). The data revealed novel information on the metabolic transition from immature to ripe fruit, but failed to consider the investigation of volatile compounds that contribute to fruit flavour formation.

The purpose of the present study was, therefore, to investigate the changes in metabolic composition from immature strawberry to ripe strawberry fruits that result in the formation of taste and nutritional quality. The biochemical changes were determined by gas chromatography–mass spectrometry (GC-MS) using either non-targeted or targeted quantitative profiling of both non-polar and polar extracts. Fruit volatile components were detected by non-targeted metabolome analysis in order to investigate fruit flavour formation. These data were complemented with high-performance liquid chromatography (HPLC) analyses of amino acids and anthocyanins. All the data underwent a variety of chemometric analyses, including principal component analysis (PCA) and partial least squares discriminant analysis (PLS-DA), to identify the true differences between samples. Both metabolic correlation analysis and metabolic network analysis based on pairwise coefficients were carried out to characterize the physiological response to developmental changes, to explore metabolic pathways, and to find bottlenecks or metabolic shifts in pathways during strawberry fruit development and ripening.

Materials and methods

Plant material and sample collection

Strawberry (*F.×ananassa* Duch. cv. Troyonoka) plants were grown in a greenhouse in the Strawberry Germplasm Resource Garden of Anhui Agricultural University with a diurnal rhythm of 16 h light and 8 h darkness following normal cultivation practices. Fruit samples were collected from 4–5 individual plants beginning with the small green fruit stage [~10 days post-anthesis (DPA)]. Between 10 and 18 fruits of the same maturation degree were harvested every 5 d, with sample collection ending at the over-ripe stage (5 d after the red-ripening stage). These fruit samples coincided with the fruit ripening stages of small green fruit (stage 1), large green fruit (stage 2), green–white fruit (stage 3), white fruit (stage 4), red-turning fruit (stage 5), red-ripening fruit (stage 6), and over-ripening fruit (stage 7) as shown in Fig. 1. After harvest, all the fruits from different ripening stages were washed in water, cut into quarters, pooled, and immediately frozen in liquid nitrogen and kept at –80 °C until use for sample preparation.

Chemicals and reagents

The standard samples of sugars and organic acids were of analytical reagent grade. Pyridine and cyclohexane were doubly distilled before use. Methoxamine hydrochloride was from the Institute of Biological Drugs Examination (Beijing, China). The reagent for trimethylsilylation, *N,O*-bis (trimethylsilyl)-trifluoroacetamide (BSTFA) containing 1% trimethylchlorosilane (TMCS), was purchased in 1 ml ampoules from Anpel Scientific Instrument Co., Ltd (Shanghai, China). A Waters ACCQ Fluor kit for HPLC analyses of amino acids was purchased from Waters Chemical Industry Co. (Milford, MA, USA). Authentic reference compounds of six anthocyanins, 3-*O*-β-glucopyranosides of delphindin (Dp3glc), 3-*O*-β-glucopyranosides of cyanidin (Cy3glc), 3-*O*-β-glucopyranosides of petunidin (Pt3glc), 3-*O*-β-glucopyranosides of pelargonidi

(Pg3glc), 3-*O*-β-glucopyranosides of peonidin (Pn3glc), and 3-*O*-β-glucopyranosides of malvidin (Mv3glc), were kindly supplied as a gift from Dr Jianhui Cheng (Institute of Horticulture, Zhejiang Agricultural Academy, Hangzhou, China) and obtained from Polyphenols Lab (Sandnes, Norway). Methanol, acetonitrile, and trichloroacetic acid were all from TEDIA (Fairfield, OH, USA) and were of HPLC grade. Water was doubly distilled and further prepared with a Purelab Classic UF purification system from Pall Co. (New York, USA).

Sample preparation

Each of the frozen fruit samples (~0.5 g) from the different ripening stages was ground into fine powder in liquid nitrogen. Each sample was mixed with 2 ml of trichloroacetic acid (0.6 mg ml^{-1}) for sample quenching and then extracted three times with 100 ml of cold (–20 °C) petroleum ether for 2 h at 4 °C. The organic phase was evaporated under nitrogen and concentrated to 1 ml. After extraction with petroleum ether, the residue was successively extracted with cold cyclohexane following the same procedure as above. Both the petroleum ether extract and the cyclohexane extract were kept with anhydrous sodium sulphate to eliminate water before GC-MS analysis. All the samples were prepared in triplicate.

The polar phase of the fruit sample was centrifuged at 12 000 *g* at 4 °C for 10 min. The fruit debris was extracted twice with 3 ml of pre-cooled methanol/water solution [1:1 (v/v) –20 °C]. The supernatant was pooled and adjusted to 5 ml with a suitable volume of the methanol/water solution. To correct for minor variations occurring during sample preparation and analysis, ribitol (0.2 mg ml^{-1} in water) was used as an internal standard. A 100 μl aliquot of the polar extract containing all hydrophilic metabolites was lyophilized under low temperature (–60 °C) in an LGJ-12 lyophilizer (Songyuan Huaxing Scientific Co., Beijing, China). Three analytical replicates were prepared for each sample. Two-step chemical derivatization was performed on the extracted hydrophilic metabolites according to the protocol of Roessner *et al.* (2000) with a few modifications. Briefly, oximation was carried out by dissolving the samples in 100 μl of methoxamine hydrochloride (25 mg ml^{-1} in pyridine) and incubating at 50 °C for 30 min. Samples were further derivatized with the addition of BSTFA containing 1% TMCS (100 μl) at 60 °C for 30 min to trimethylsilylate the polar functional groups. The derived samples were equilibrated to room temperature before injection.

GC-MS analysis

Capillary GC-MS analysis of the petroleum ether extract and the cyclohexane extract was performed with a SHIMADZU QP2010 series (Shimadzu Instruments, Japan) coupled to a Fisons MD800 quadrupole mass detector (Fisons Instruments, CA, USA) fitted with a split injector (1:10). Each sample (1 μl) was injected into the gas chromatograph through a fused-silica capillary column (30 m×0.25 mm i.d., 0.25 μm DB-5 MS stationary phase, J&W Scientific, Folsom, CA, USA). It was operated in constant pressure mode at 91 kPa and the injector temperature was 250 °C. The column temperature was held at 100 °C for 1 min; increased to 150 °C with a temperature gradient of 20 °C min^{-1}; increased to 200 °C at 4 °C min^{-1}; increased to 280 °C at 10 °C min^{-1}; then held for 10 min. Helium (99.999%) was used as the carrier gas with the flow rate at 1 ml min^{-1}. The significant MS operating parameters were as follows: ionization voltage was 70 eV (electron impact ionization), ion source temperature was 200 °C, and interface temperature was 250 °C. Ions were generated by a 70 eV electron beam at an ionization current of 40 μA. TIC (total ion current) spectra were recorded in the mass range of 40–600 atomic mass units (amu) in scanning mode. Shimadzu GCMS solution software (version 2.30) was used for data acquisition.

Polar extract analysis by GC-MS was carried out using the same equipment as above, but with modifications in the GC program.
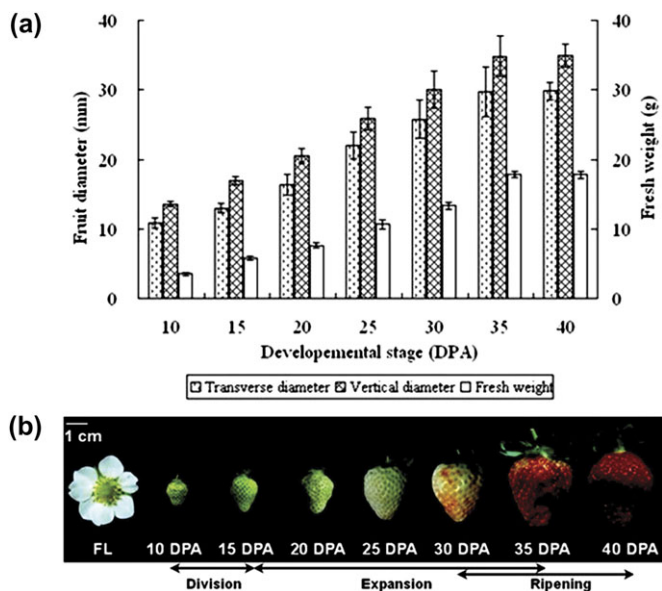
Fig. 1. Growth and development of 'Troyonoka' strawberry fruits cultivated in a glass greenhouse. (a) Changes in fruit transverse diameter, vertical diameter, and fruit fresh weight. Mean of 10–18 fruits. Vertical bars represent the SD. (b) Stages of fruit development from flower (FL) to over-ripening fruit and the corresponding days after anthesis (DPA) separated into three phases of development characterized by cell division, cell expansion, and fruit ripening.

The GC was programmed at an initial temperature of 100 °C for 2 min; increased to 184 °C at 3 °C min⁻¹; increased to 200 °C at 0.5 °C min⁻¹; held for 3 min; increased to 280 °C at 15 °C min⁻¹; and finally held for 10 min. The temperatures of the injection port (splitless mode), interface, and MS source were 250, 250, and 200 °C, respectively.

HPLC analysis on amino acids and anthocyanins

For amino acid analysis, a pre-column derivatization method was used for sample preparation. Frozen fruit slices were homogenized in 20 µl of boiling water per mg of fresh weight of tissues and extracted at 100 °C for 30 min. After centrifugation at 11 000 *g* for 5 min, the fruit debris was discarded, and the extracts were immediately used for derivatization. A 10 µl aliquot of supernatant was transferred and mixed successively with 70 µl of borax buffer (pH 8.0) and 20 µl of ACCQ Fluor derivatization reagent (6-aminoquinolyl-*N*-hydroxyl-succinimidyl carbamate). The mixture was kept at 55 °C for 10 min and then subjected to HPLC analysis. The chromatographic analysis was performed on a C₁₈ reverse phase column (Nova-Pak™, 250 mm×4.6 mm, particle size 5 µm) at 37 °C. The HPLC system was a Waters 600E (Milford, MA, USA) equipped with a binary 2489 UV detector and a 2475 fluorescence detector. The excitation and emission wavelengths of the fluorophotomeric detector were 250 nm and 395 nm, respectively, and the UV detector wavelength was set at 248 nm. The elution gradient consisted of acetate phosphate buffer (A), acetonitrile (B), and water (C), and the elution profile was: 0 min, 100% A; 18 min, 5% B; 19 min, 9% B; 29.5 min, 17% B; 33 min, 60% B, 40% C; 36 min, 100% A; 45 min, 60% B using linear gradients in between the time points. A 5 µl aliquot of each fruit sample was manually injected and the flow rate was maintained at 1 ml min⁻¹. All data were processed using Waters Empower Station.

For anthocyanin analysis, fine powders of each fruit sample (0.5 g) were extracted for 20 min with 5 ml of extraction buffer (acetone:water:formic acid 80:19.8:0.2) in the dark. The mixture was centrifuged at 11 000 *g* for 10 min and adjusted to 5 ml with extraction buffer. The supernatant was further filtered through a syringe-tip cellulose acetate filter (0.45 µm) prior to use. HPLC was performed on a Waters 600E system equipped with a 2487 UV detector. Chromatographic separation was carried out on a (250 mm×4.6 mm) Hyersil ODS 5 µm C₁₈ HPLC column (Supelco Inc., Bellefonte, PA, USA). The stationary phase consisted of 5% formic acid in water (A), while the mobile phase was acetonitrile (B). The gradient program was linear from 0% to 20% B (0–13 min); 20–30% B (13–20 min); and 30–0% B (20–25 min). Prior to the next injection, the column was equilibrated for 30 min with 5% formic acid. The column temperature was maintained at 25 °C and the flow rate was maintained at 1 ml min⁻¹. The chromatographic profile was recorded at 520 nm. All data were processed using Millinum³² software (Waters Co., MA, USA).

Identification of metabolites detected by GC-MS and HPLC

Recorded data profiles from GC-MS analysis were converted into NetCDF format and then exported to the MetAlign software package (RIKILT-WUR Institute of Food Safety, Wageningen, The Netherlands), where all data pre-treatment procedures including noise reduction, baseline correction, and peak alignment were carried out. Artefact peaks, such as column bleeding, and solvent peaks were removed. For metabolite identification and annotation, peaks were matched against customized reference spectrum databases including the National Institute of Standards and Technology (NIST) and the Wiley Registry, based on retention indices and mass spectral similarities. Some constituents, including the major sugars and organic acids, were further confirmed by comparing their mass spectra and retention indices with those of authentic reference compounds (Supplementary Figs S1, S2 available at *JXB* online). Metabolite contents were

calculated by the internal standard method and independently normalized to plant mg of fresh weight and to internal references (*n*-heptadecane in the non-polar phase and ribitol in the polar phases) with each sample collection point.

The chromatographic peaks in the polar extracts were identified and quantified by comparing their retention times and UV visible absorption spectra with those of authentic reference compounds. The amounts of each amino acid and anthocyanin were calculated using regression equations.

The data presented in Supplementary Table S1 at *JXB* online are the average of the three replicate samples. The ratio of metabolite content at each sample collection point to the average content of each metabolite across seven developmental stages was shown by the colour scale. The heat map analysis was done on a Microsoft Excel-implemented macro program available at http://bitesizebio.com/2009/02/03/how-to-create-a-heatmap-in-excel.

Multivariate analysis

For individual metabolites, means ±SD were calculated from three biological replicates. Mean comparisons between the harvest stages were conducted by a Duncan's *t*-test using SPSS software version 13.0 (SPSS Institute, Chicago, IL, USA). A difference with $P < 0.05$ was considered significant. The ratios calculated by the means of the metabolite contents were used to explain the up-regulation or down-regulation of identified metabolites.

The GC-MS data matrix after peak annotation and normalization was used for multivariate analysis with SIMCA-P 11.5 software (Umetrics AB, Umeå, Sweden). PCA was used for an unsupervised analysis and PLS-DA for a supervised analysis. A PCA scatter plot of variables showing different metabolites was obtained to explain the separation of polar phase samples from different strawberry fruit stages. A range of metabolites was selected as the variable importance in the projection (VIP) based on PCA.

Metabolite correlation and network analysis

To visualize the correlation matrix, pairwise metabolite correlations were calculated by Pearson's correlation coefficient (r_{ij}) within and between 224 annotated peaks in the non-polar phase and 107 annotated peaks and 17 amino acids in the polar phase. The level of significance was set as $|r_{Met}| \geqslant 0.85$ for both the non-polar and polar phase. An $|r_{Met}| \geqslant 0.80$ was used for the correlation analysis between the non-polar phase and the polar phase. Correlation networks were reconstructed from the measured metabolite concentrations by pairwise combination of the metabolite concentrations, and 1111 pairwise analyses with strong dependencies (correlation coefficients: $|r_{Met}| \geqslant 0.85$) were selected and applied for metabolic network analysis. The network of pairwise metabolite correlation is drawn using the Kamada kawei-algorithm in Pajek (http://vlado.fmf.uni-lj.si/pub/networks/pajek/). In the correlation network, there is an edge between two metabolites if the absolute value of the correlation coefficient resulting from two concentration vectors exceeds a certain threshold $|r_{Met}|$, and all the scattered vertexes are deleted. Publicly available data from the KEGG pathway database were obtained to understand and confirm the relationship between metabolite–metabolite correlations.
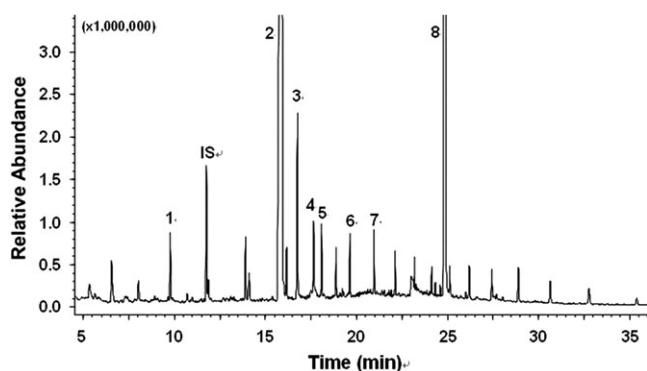
Results

Fruit metabolite screening by GC-MS

Strawberry was grown in a greenhouse and different stages of fruits were harvested for metabolite analysis. These fruit samples were collected every 5 d as shown in Fig. 1. Initial GC-MS analyses of the non-polar and polar metabolites

from all seven stages identified >280 resolved peaks, including the small green, large green, green–white, white, red-turning, red-ripening, and over-ripening fruit stages. Approximately 40% of these peaks could be identified as discrete metabolites with known chemical structure (Supplementary Table S1 at *JXB* online). These identified metabolites represent numerous metabolic pathways, including biosynthesis of isobutyl phthalate, isooctyl phthalate, and bis(2-ethylhexyl) phthalate, and the metabolism of related compounds [e.g. the tricarboxylic acid cycle (TCA) and the shikimate pathway]. The presence of 2,6-dimethyl-α-D-galactopyranoside and 1-methyl-α-D-galactopyranoside, which are considered possible precursors for compounds responsible for strawberry fruit aroma, was also noted. These compounds have been reported previously in flavour components of grape fruit (Mateo *et al.*, 1997) and oolong tea (Guo *et al.*, 1993).

In the petroleum ether extracts of red-ripening fruits, 21 components from 35 chromatographic peaks were identified. Through peak area integration, the relative amounts of isobutyl phthalate, isooctyl phthalate, 1,2-benzenedicarboxylic-bis (2-methylpropyl) ester, and dibutyl phthalate were 856.44, 1336.54, 269.71, and 168.89 µg g^{-1}, respectively. Around 615.76 µg g^{-1} of fatty acids and alcohols were represented by *n*-hexadecanoic acid; 9,12-(*Z, Z*)-octadecadienoic acid; 3,7,11-trimethyl-1,6,10-dodecatrien-3-ol; and 9,12,15-(*Z, Z, Z*)-octadecatrien-1-ol, which have an important influence on the flavour of strawberry fruit (Fig. 2; Supplementary Table S1 at *JXB* online).

In the cyclohexane extract of red-ripening strawberry fruits, 107 volatile metabolites from 160 chromatographic peaks were identified using the NIST and Wiley databases. The most relatively abundant metabolites were alkanes, esters, and alcohols, including hexadecane (643.36 µg g^{-1}), *n*-tetradecane (52.23 µg g^{-1}), nonadecane (49.20 µg g^{-1}), *n*-docosane (69.77 µg g^{-1}), *n*-tricosane (29.71 µg g^{-1}), *n*-octadecane (48.01 µg g^{-1}), undecane (39.63 µg g^{-1}), eicosane (25.59 µg g^{-1}), *n*-dodecane (11.20 µg g^{-1}), isobutyl

phthalate (453.08 µg g^{-1}), bis(2-ethylhexyl) phthalate (716.13 µg g^{-1}), 1,2-benzenedicarboxylic acid, butyl 8-methylnonyl ester (69.48 µg g^{-1}), 5-methyl-1-heptanol (118.98 µg g^{-1}), and 8-dodecen-1-ol (5.83 µg g^{-1}) (Fig. 3; Supplementary Table S1 at *JXB* online). Comparing compounds identified in this study with the perceived aroma characteristics of different volatile compounds reveals that strawberry aroma is the result of the combined perception of fruity, green, sweaty, peach-like, and caramel-like flavour (Larsen and Poll, 1992).

Major metabolites in the polar metabolic extracts of strawberry fruits are sugars, glucosides, sugar alcohols, and organic acids (Fig. 4), which showed significant differences in levels during fruit growth and ripening. For quantifications, oximation and derivatization treatments were applied. GC-MS profiling analysis on red-ripening strawberry fruits showed that the average contents of fructose,
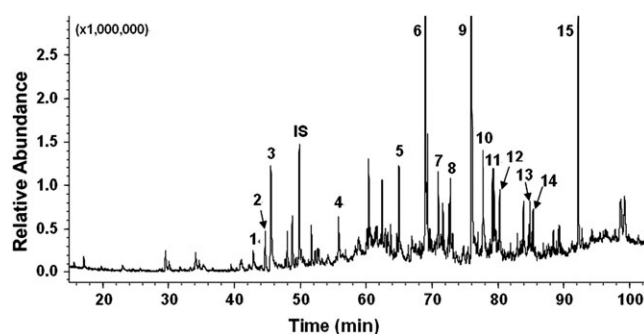
Fig. 3. GC-MS analysis on cyclohexane extract from strawberry fruits (TIC). (1) *n*-Dodecane; (2) 8-dodecen-1-ol; (3) 5-methyl-1-heptanol; (4) 2,6,10,14-tetramethyl-heptadecane; (5) *n*-tetradecane; (6) hexadecane; (7) *n*-tricosane; (8) nonadecane; (9) isobutyl phthalate; (10) *n*-docosane; (11) 1,2-benzenedicarboxylic acid, butyl 8-methylnonyl ester; (12) *n*-octadecane; (13) eicosane; (14) undecane; (15) bis(2-ethylhexyl) phthalate; IS, internal standard.
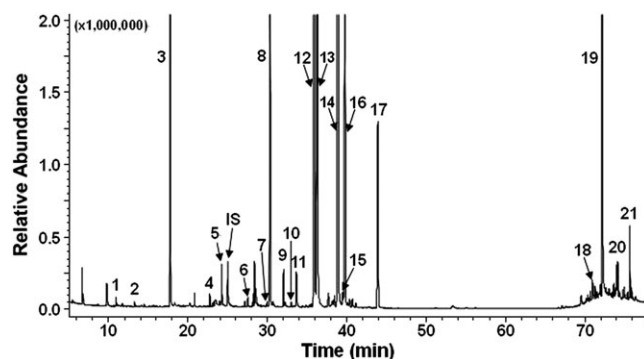
Fig. 2. GC-MS analysis on petroleum ether extract from strawberry fruits (TIC). Labelled peaks are: (1) 3,7,11-trimethyl-1,6,10-dodecatrien-3-ol; (2) isobutyl phthalate; (3) 1,2-benzenedicarboxylic-bis(2-methylpropyl) ester; (4) *n*-hexadecanoic acid; (5) dibutyl phthalate; (6) 9,12-(*Z, Z*)-octadecadienoic acid; (7) 9,12,15-(*Z, Z, Z*)-octadecatrien-1-ol; (8) isooctyl phthalate; IS, internal standard.

Fig. 4. GC-MS analysis on polar metabolites from strawberry fruits (TIC). (1) Butanedioic acid; (2) 2,5-dimethyl-4-hydroxy-3(2H)-furanone; (3) malic acid; (4) arabinonic acid; (5) gluconic acid; (6) mannonic acid; (7) 2-keto-D-gluconic acid; (8) citric acid; (9) 2-hydroxyethylsulphonic acid; (10) gulonic acid; (11) glucose; (12 and 13) fructose; (14 and 16) galactose; (15) palmitic acid; (17) *myo*-inositol-1-phosphate; (18) hexadecanoic acid; (19) sucrose; (20) turanose; (21) palatinose; IS, internal standard.

sucrose, galactose, and turanose were 14.76, 9.80, 7.75, and 1.87 mg g^{-1}, respectively, while those of glucose and palatinose were 210.42 µg g^{-1} and 190.17 µg g^{-1}, respectively. Organic acids, primarily comprised of citric acid and malic acid, were 7.74 mg g^{-1} and 4.22 mg g^{-1}, respectively, while other compounds (e.g. 2-keto-D-gluconic acid, palmitic acid, *n*-hexadecanoic acid, arabinonic acid, butanedioic acid, gulonic acid, gluconic acid, mannonic acid, and 2-hydroxy-ethyl-sulphonic acid) made up ~2% of the identified polar metabolites. *Myo*-inositol-1-phosphate was at 1.37 mg g^{-1}.

Untargeted profiling of metabolites by GC-MS

The parallel analysis of non-polar and polar extracts using GC-MS enabled the detection of a large number of compounds of different classes, mainly alkanes, alcohols, esters, sugars and sugar alcohols, organic acids, and fatty acids (Supplementary Table S1 at *JXB* online). The heat map showed that the metabolites have a differential distribution during development and maturation. Almost all alkanes exhibited significance increases during fruit development, while the accumulation pattern of each metabolite was very different. Some alkanes (e.g. tridecane, tetradecane, pentadecane, 2,6,10-trimethyltetradecane, octadecane, 2,6,10,14-tetramethyl-hexadecane, and nonadecane) began to accumulate at the red-turning stage, and were maintained at high levels; others increased through stages 4–6 initially but decreased thereafter (Fig. 5). Most of the alcohols and esters reached the highest levels in red-turning and red-ripening fruits, and these contributed to the formation of fruity, sweaty, and peach-like flavour in strawberry fruits (Larsen and Poll, 1992). Sugars, organic acids, and fatty acids were the metabolites that exhibited a high degree of variance during fruit development (Fig. 5; Supplementary Table S1 at *JXB* online), while fruit ripening was characterized by the increases in levels of fructose, sucrose, galactose, citric acid, and malic acid, which comprised the major soluble sugars and organic acids in strawberry fruits.

Cluster analysis allowed the distinction between samples that were collected at the different stages. Important metabolites in these samples that affected cluster formation could be determined. Multivariate statistical analysis PCA and PLS-DA were applied to non-polar and polar extracts separately. Table 1 lists the top 20 metabolites that influenced cluster formation within the non-polar metabolite profiles, which were generated from PLS-DA analysis of the cyclohexane extract. These compounds include alkanes, aldehydes, and esters (Table 1 and Fig. 7). A different cluster pattern from the polar metabolite extracts was observed through PLS-DA analysis. Specifically, five metabolite profiles from stage 3 to stage 7 fruit samples clearly separated from each other, forming distinct clusters (Fig. 6). Stage 1 and stage 2 fruit profiles (both green fruit) clustered together and were clearly separated from the five ripening profiles (Fig. 6).

PCA uses an *n*-dimensional vector approach to separate samples based on the cumulative correlation of all data. The resulting vectors that yield the greatest separation between samples are identified and then used to calculate their factor scores (Roessner *et al.*, 2001). The first two highest-ranked vectors, which represent the major variance among samples, were then displayed in two-dimensional plots. A broad range of metabolites were found to influence cluster formation of polar metabolite profiles, including sugars, glycosides, sugar alcohols, organic acids, and 2-deoxyery-thropentono-1,4-lactone (Supplementary Fig. S3 at *JXB* online; Table 2). The accumulation profiles of VIPs in the polar phase showed that strawberries underwent major changes in carbohydrate concentrations during fruit development (Fig. 7).

Targeted profiling of amino acids and anthocyanins by HPLC

Free amino acids not only participate in a wide range of physiological reactions, especially in carbon–nitrogen metabolism in fruits, but also contribute to fruit quality formation. HPLC coupled with fluorescence detection was used to investigate the changes of the main amino acid components during the development of strawberry fruits. The amino acid components in strawberry fruits can be well separated (Supplementary Fig. S4 at *JXB* online). Serine, arginine, glutamic acid, histidine, aspartic acid, and proline were all present at significantly high levels in strawberry fruits (Supplementary Table S1). PLS-DA analysis on the profiles of amino acids gave a clear cluster pattern (Fig. 8). During the development of strawberry fruits, the total content of amino acids decreased gradually, from 4911.91 µg g^{-1} in stage 1 to 894.04 µg g^{-1} fresh weight in stage 6 (Supplementary Table S1). During stage 1 to stage 6, the level of each amino acid component decreased (Supplementary Table S1; Fig. 5). However, the levels of most amino acids showed a significant increase in stage 7 when compared with stage 6, which might be related to cell degradation in over-ripening fruits (Fig. 9B).

In previous studies, 25 anthocyanin pigments were detected from five different strawberry cultivars, with most cultivars containing pelargonidin (Pg) and cyanidin (Cy) (Lopes *et al.*, 2007). It was found that the major anthocyanin in the ripening fruits of cultivar 'Troyonoka' is Pg3glc, which began to accumulate in red-turning fruits (Supplementary Table S1 at *JXB* online). A significant increase (30.4-fold) was observed in red-ripening fruits (stage 6). The level of Pg3glc continued to increase (1.2-fold) in over-ripening fruits (stage 7) as compared with stage 6 (Fig. 9). The smaller proportions of Cy3glc varied drastically from stage 5 to stage 7. Cy3glc showed a large increase (99.4-fold) between stage 5 and stage 6, then decreased significantly (0.5-fold) by stage 7 (as compared with stage 6). Pn3glc and Mv3glc were detected only in red-ripening fruits in a minute quantity (Supplementary Table S1). Notable variability was found among the anthocyanin concentrations in fruit samples at different harvest times, indicating a strong influence from the degree of maturity of the strawberry fruits on organoleptic quality.
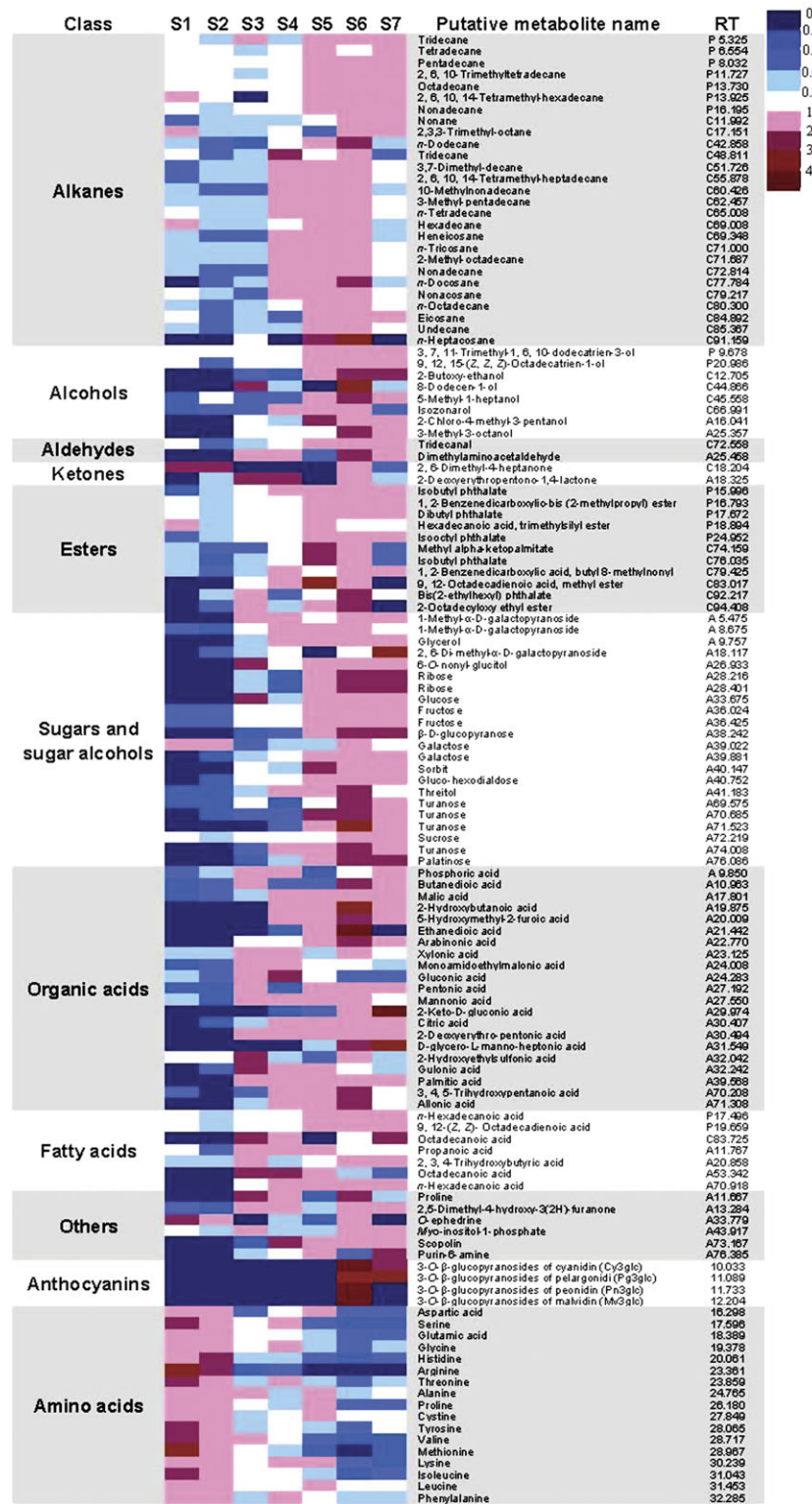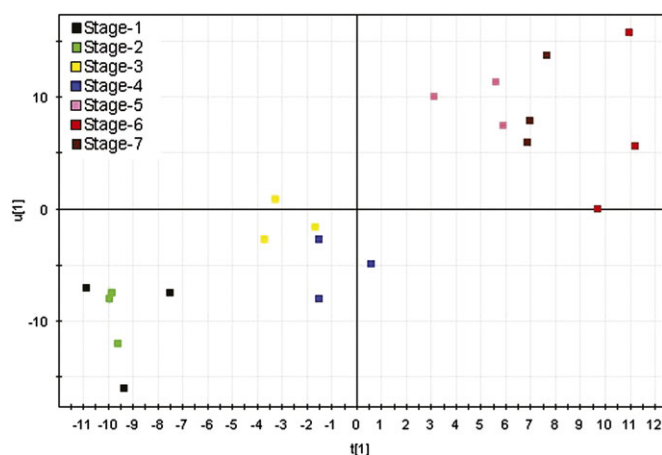
Fig. 5. Metabolite levels of representatives of major metabolite groups which changed significantly during fruit development. The amounts were normalized to the fresh weight of each developmental stage (μg g^{-1}) and are provided in Supplementary Table S1 at *JXB* online. The ratio of metabolite content at each sample collection point to the average metabolite content across seven developmental stages is shown by the colour scale. The lowest ratios are in dark blue, and the highest ratio in dark red. S1–S7 refer to the seven developmental stages of strawberry fruits shown in Fig. 1.

Table 1. Compounds determined to be of variable importance in the projection through PLS-DA analysis on the identified compounds from the cyclohexane extract

No.	RT (min)	Name	VIP
1	65.008	*n*-Tetradecane	1.723
2	72.558	Tridecanal	1.671
3	42.858	*n*-Dodecane	1.663
4	79.217	Nonacosane	1.660
5	71.687	2-Methyl-octadecane	1.650
6	69.008	Hexadecane	1.618
7	79.425	1,2-Benzenedicarboxylic acid, butyl 8-methylnonyl ester	1.617
8	66.991	Isozonarol	1.612
9	72.814	Nonadecane	1.598
10	77.784	*n*-Docosane	1.576
11	71.000	*n*-Tricosane	1.541
12	84.892	Eicosane	1.537
13	94.408	2-Octadecyloxy ethyl ester	1.531
14	74.159	Methyl α-ketopalmitate	1.516
15	91.159	*n*-Heptacosane	1.513
16	80.300	*n*-Octadecane	1.512
17	92.217	Bis(2-ethylhexyl) phthalate	1.462
18	85.367	Undecane	1.451
19	69.348	Heneicosane	1.435
20	76.035	Isobutyl phthalate	1.409

RT, retention time; VIP, variable importance in the projection.

Table 2. Compounds determined to be of variable importance in the projection through PLS-DA analysis on the identified compounds from the polar phase extract

No.	RT (min)	Name	VIP
1	18.325	2-Deoxyerythropentono-1,4-lactone	1.799
2	21.442	Ethanedioic acid	1.668
3	24.283	Gluconic acid	1.609
4	71.523	Turanose	1.578
5	8.675	1-Methyl-α-D-galactopyranoside	1.569
6	72.219	Sucrose	1.557
7	38.242	β-D-Glucopyranose	1.539
8	70.685	Turanose	1.528
9	27.550	Mannonic acid	1.468
10	24.008	Monoamidoethylmalonic acid	1.444
11	5.475	1-Methyl-α-D-galactopyranoside	1.433
12	18.117	2,6-di-Methyl-α-D-galactopyranoside	1.431
13	27.192	Pentonic acid	1.294
14	40.752	Glucohexodialdose	1.293
15	41.183	Threitol	1.281
16	73.167	Scopolin	1.269
17	36.425	Fructose	1.252
18	29.817	Butanoic acid	1.223
19	40.147	Sorbit	1.196
20	69.575	Turanose	1.185

RT, retention time; VIP, variable importance in the projection.

Fig. 6. PLS-DA score plot of polar phase samples from different strawberry fruit stages. The coloured dots represent polar phase samples from different strawberry fruit stages. Seven independent time course studies, each with triplicate samples taken for extraction at each time point, were used in the analysis, for a total of 21 data points.

Metabolite changes in red-turning and over-ripening stages of strawberry fruits

In order to identify the metabolic pathways that modulate metabolite levels linked to peak market quality, significant metabolite changes (P <0.05) observed between either stage 5 or stage 7 and stage 6 were highlighted on the metabolic map (Fig. 9). In stage 6, significant increases in fructose (1.2-fold), sucrose (1.3-fold), galactose (1.5-fold),

citric acid (1.5-fold), and malic acid (1.5-fold), which are the main polar metabolites contributing to quality, were found. Up-regulation was also observed in maltose (5.5-fold), glucose (1.6-fold), *myo*-inositol-1-phosphate (1.9-fold), and glucose-6-phosphate (1.5-fold) (Fig. 9A). In contrast, there was a significant decrease in the levels of glycerol (0.7-fold), rhamnose (0.8-fold), erythrose (0.7-fold), gluconic acid (0.6-fold), and sorbitol (0.5-fold). The levels of most of the amino acids also decreased from stage 5 to stage 6.

The development through stage 6 into stage 7 results in an almost complete metabolic reversal. The levels of major sugars and organic acids, such as fructose (0.9-fold), galactose (0.8-fold), citric acid (0.7-fold), and malic acid (0.7-fold), showed significant decreases (Fig. 9B). In addition, *myo*-inositol-1-phosphate (0.7-fold), *myo*-inositol (0.7-fold), and glycerol-3-phosphate (0.6-fold) were sharply decreased in stage 7. The levels of sucrose and some amino acids, such as serine, cysteine, leucine, proline, aspartate, arginine, and lysine, remain the same. As shown above, most other amino acids showed a significant increase in stage 7, which suggests that cell degradation occurs in over-ripening fruits (Knee *et al.*, 1977).

Correlation of metabolite levels and analysis of dependencies of non-polar metabolites

Correlation analysis is a useful tool to explore metabolic pathways and networks (Raamsdonk *et al.*, 2001; Roessner *et al.*, 2001; Fiehn, 2003; Steuer *et al.*, 2003a, b; Weckwerth *et al.*, 2004). In order to delineate metabolic changes relevant to fruit ripening, the pairwise correlation for each
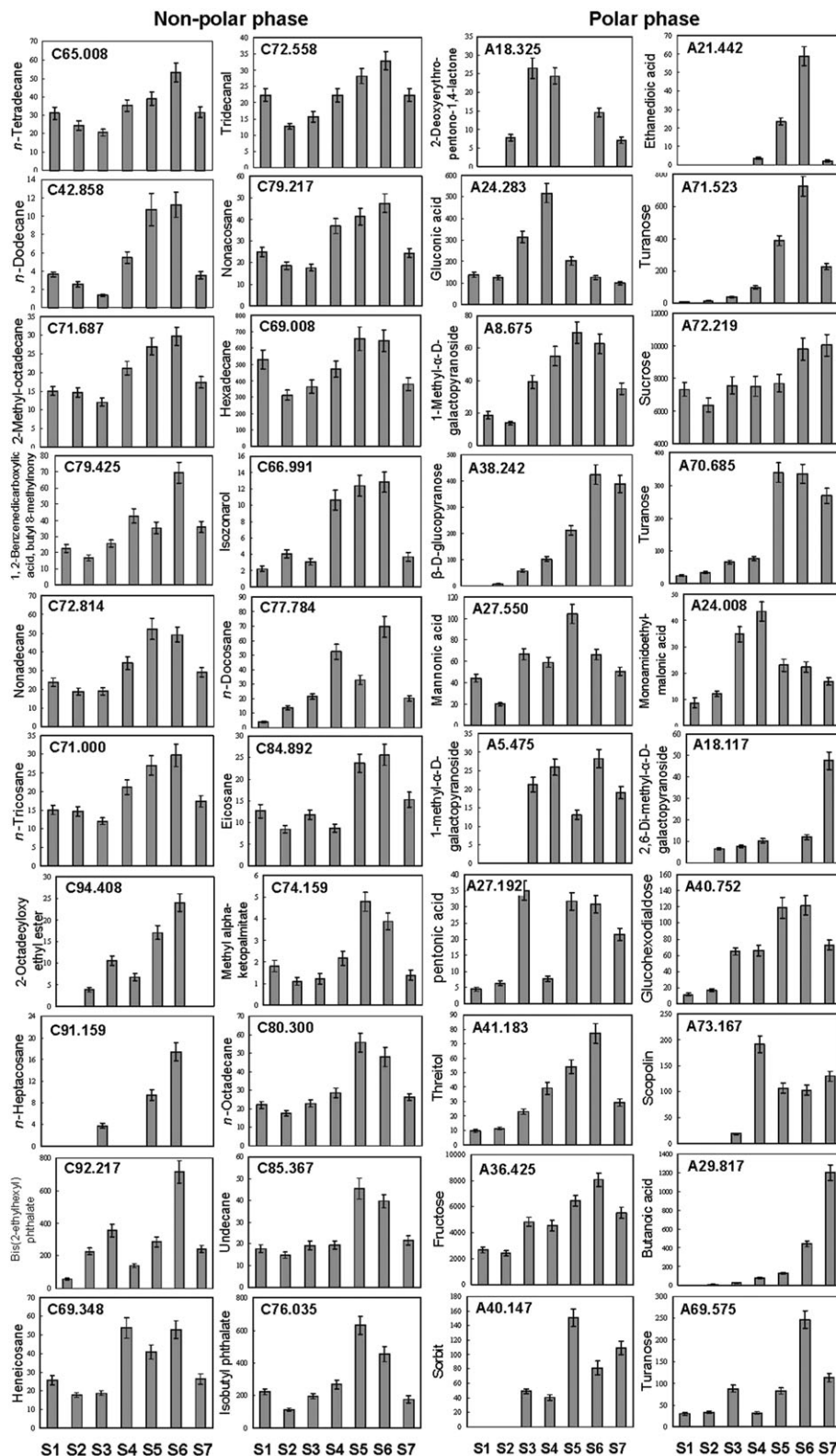
Fig. 7. Key compounds identified from the PLS-DA (VIPs) showed major differences over fruit maturation. Metabolites from the non-polar phase and polar phase were detected by GC-MS. The bars represent metabolite contents across the seven different developmental stages (S1–S7) of strawberry fruits. The amounts are normalized to the fresh weight of each developmental stage (μg g^{-1}). The error bars are the standard deviations from three biological repeats.
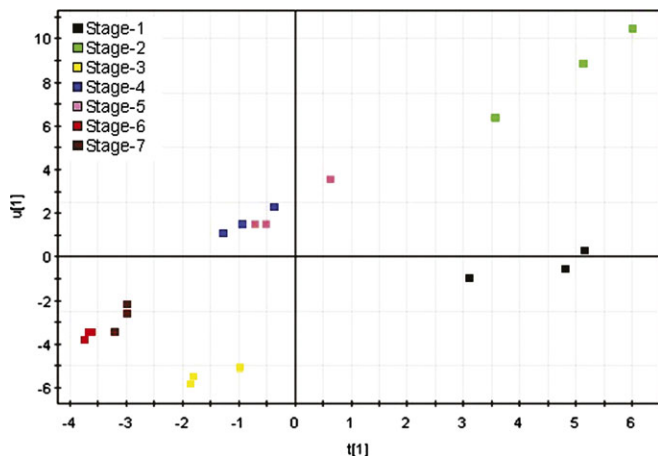
Fig. 8. PLS-DA score plot of amino acid extracts from different strawberry fruit stages. The coloured dots represent polar phase samples from different strawberry fruit stages. Seven independent time course studies, each with triplicate samples taken for extraction at each time point, were used in the analysis, for a total of 21 data points.

metabolite at different stages was analysed against every other metabolite within the samples.

For non-polar compounds, strong correlations ($|r_{Met}| \geqslant 0.8$) were found in three main metabolites [hexadecane, isobutyl phthalate, and bis(2-ethylhexyl) phthalate] from the cyclohexane extract (Supplementary Table S2 at *JXB* online). These relationships could be classified into three types: (i) those between compounds having similar chemical structures or belonging to the same chemical class, such as alkanes (e.g. *n*-dodecane, *n*-etradecane, hexadecane, non-adecane, and octadecane), and esters (e.g. isobutyl phthalate, 1,2-benzenedicarboxylic acid, and butyl 8-methylnonyl ester); (ii) those between compounds having different chemical structures, but that are known to be connected by close biosynthetic relationships (e.g. 2-ethylhexyl phthalate and 2-ethyl-1-hexanol, because 2-ethyl-1-hexanol is used as the precursor in the synthesis of 2-ethylhexyl phthalate); and (iii) those that seem unrelated based on available knowledge (e.g. isobutyl phthalate and alkanes) (Supplementary Table S2).

Correlation analysis of polar metabolites

In the polar extract of strawberry fruits, 107 polar metabolites identified from untargeted metabolic profiling analysis by GC-MS and 17 amino acids from targeted metabolic profiling analysis by HPLC were subjected to correlation analysis. In total 7626 pairwise analyses and their Pearson correlation coefficients (r_{ij}) were computed. Metabolites with correlation coefficients ($|r_{Met}| \geqslant 0.85$) were set as strong dependencies and 1111 pairwise analyses were selected (14.57% of total correlations). Overall, both positive and negative dependencies were observed; the most significant correlations were positive among these polar metabolites, albeit few in number. Most of the coefficients were negative.

The correlation pairs of polar metabolites were categorized into similar groups as described above with respect to biosynthetic relationships: (i) metabolite pairs placed close to each other in a metabolic pathway for the concerned reactions, for example the pair of sucrose:fructose ($r=0.944$, $P=0.14\%$), sucrose:glucose ($r=0.956$, $P=0.08\%$), citric acid:-malic acid (in the TCA cycle, $r=0.989$, $P=0.02\%$), and phenylalanine:tyrosine (in the shikimate pathway, $r=0.951$, $P=0.10\%$); (ii) metabolite pairs with a similar chemical skeleton derived from the same biosynthetic pathway, for example the pairs galactose:1-methyl-α-D-galactopyranoside ($r=0.852$, $P=1.48\%$) and glucose:glucopyranoside ($r=0.866$, $P=1.17\%$); and (iii) metabolite pairs located at a distance in a metabolic network, for example the pairs serine:glutamate ($r=0.956$, $P=0.08\%$) and glutamic acid:arginine ($r=0.980$, $P=0.01\%$) in arginine metabolism (Supplementary Fig. S5 at *JXB* online).

Correlation analysis between non-polar metabolites and polar metabolites

Although the non-polar metabolites and polar metabolites were measured in two separate chromatographic analyses for any given time point, they came from the same tissue sample. It is therefore possible to make connections between these profiles by correlation analysis. Pairwise metabolite correlations between 224 non-polar metabolites and 124 polar metabolites (107 annotated peaks and 17 amino acids in the polar phase) were calculated by Pearson's correlation coefficient, with the threshold set as $|r_{Met}| \geqslant 0.80$. It was found that isobutyl phthalate showed strong positive correlations with mannonic acid ($r=0.885$, $P=0.08\%$) and glucohexodialdose ($r=0.842$, $P=1.74\%$). Bis(2-ethylhexyl) phthalate exhibited strong positive correlations with 2,3,4-trihydroxybutyric acid ($r=0.897$, $P=0.62\%$), turanose ($r=0.886$, $P=0.79\%$), sucrose ($r=0.927$, $P=0.27\%$), and palatinose ($r=0.849$, $P=1.56\%$). 5-Methyl-1-heptanol strongly correlated with ethanedioic acid ($r=0.937$, $P=0.18\%$), fructose ($r=0.845$, $P=1.66\%$), turanose ($r=0.900$, $P=0.57\%$), sucrose ($r=0.991$, $P=0.01\%$), and palatinose ($r=0.972$, $P=0.01\%$). Fructose, galactose, and malic acid showed strong correlations with 2-butyl-1-octanol, with the correlates 0.909 ($P=0.05\%$), 0.821 ($P=2.35\%$), and 0.865 ($P=1.20\%$), respectively. Fructose, malic acid, and citric acid all have positive correlations with *n*-octadecene, with the correlates 0.810 ($P=2.73\%$), 0.879 ($P=0.09\%$), and 0.883 ($P=0.08\%$), respectively. The data presented here provide the correlative information that may, with further experimentation, allow the elucidation of the linkage between the biosynthesis pathways of sugars, organic acids, alcohols, esters, and alkenes.

Metabolic network analysis

In addition to correlation analysis, network analysis based on a correlation matrix was carried out to visualize the network and focus on a few metabolites of interest. This network analysis allowed the definition of the key points at
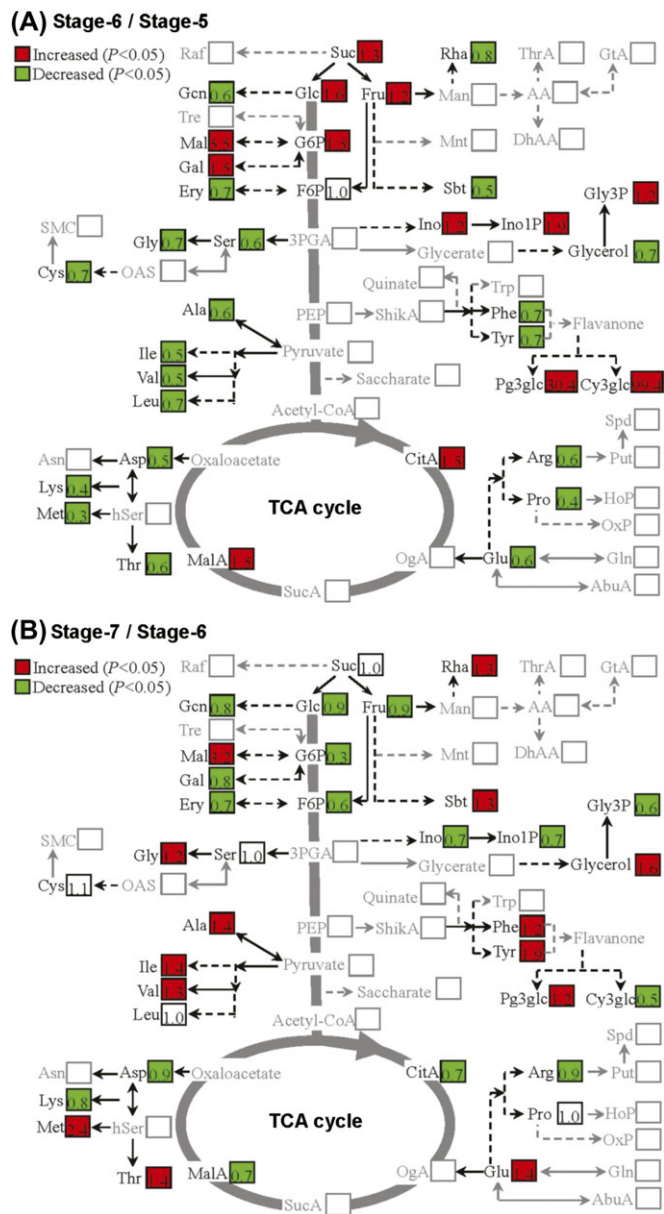
Fig. 9. Changes in the levels of metabolites during fruit development shown in a metabolic diagram. (A) The ratios between stage 6 and stage 5. (B) The ratios between stage 7 and stage 6. The changes in metabolite contents were calculated by dividing the metabolite level in fruit ripening stages 5, 6, and 7. The level of significance was set at P <0.05. The metabolites with grey characters were undetectable. 3-PGA, 3-phosphoglycerate; AA, ascorbate; AbuA, 4-aminobutyrate; Ala, alanine; Arg, arginine; Asn, asparagine; Asp, aspartate; CitA, citric acid; Cy3glc, 3-O-β-glucopyranosides of cyanidin; Cys, cysteine; DhAA, dehydroascorbate; Ery, erythrose; Fru, fructose; G6P, glucose-6-phosphate; Gal, galactose; Gcn, gluconic acid; Glc, glucose; Gln, glutamine; Glu, glutamate; Gly, glycine; Gly3P, glycerol-3-phosphate; GtA, galacturonate; HoP, hydroxyproline; hSer, homoserine; Ile, isoleucine; Ino, *myo*-inositol; Ino1P, *myo*-inositol-1-phosphate; Leu, leucine; Lys, lysine; Mal, maltose; MalA, malic acid; Man, mannose; Met, methionine; Mnt, mannitol; OAS, O-acetyl serine; OgA, α-oxoglutarate; OxP, 5-oxoproline; PEP, phosphoenolpyruvic acid; Pg3glc, 3-O-β-glucopyranosides of Pelargonidi; Phe, phenyl-alanine; Pro, proline; Put, putrescine; Raf, raffinose; Rha, rhamnose; Sbt, D-sorbitol; Ser, serine; ShikA, shikimate; SMC, S-Me-cysteine; Spd, spermidine; Suc, sucrose; SucA, succinate; Thr, threonine; ThrA, threonate; Top, α-tocopherol; Tre, trehalose; Trp, tryptophan; Tyr, tyrosine; Val, valine.

which metabolism changed in metabolic networks using the connectivity matrix (Weckwerth *et al.*, 2004).

After correlation analysis of polar metabolites and amino acids, 1111 pairwise correlations with strong dependencies (correlation coefficients: |r_{Met}| ≥0.85) were picked and applied for metabolic network analysis. Supplementary Figure S5 at *JXB* online presents the metabolite correlation network in the polar phase extract of strawberry fruits, and each vertex corresponds to a metabolite. All the scattered vertexes are deleted and the metabolites with significant correlations are linked together.

Fructose, galactose, citric acid, and malic acid, the four main polar metabolites closely related to ripe fruit, exhibited strong positive correlations with the following metabolites: sucrose, glucose, ribose, turanose, glucohexodialdose, palatinose, glucopyranoside, 1-methyl-α-D-galactopyranoside, mannonic acid, and arabinonic acid; but had negative correlations with serine, glutamic acid, threonine, and isoleucine. Moreover, fructose showed strong positive correlations with other metabolites, including erythrose, hexadecanoic acid, and scopolin, and negative correlations with glycine and histidine. Fructose was placed close to erythrose in the metabolic network because fructose 6-phosphate can be converted into erythrose 4-phosphate by transketolase. Galactose had strong positive correlations with 2-keto-D-gluconic acid, threitol, D-glycero-L-manno-heptonic acid, idonic acid, and scopolin, but had negative correlations with glycine, histidine, arginine, and methionine. Citric acid exhibited strong positive correlations with 2-hydroxybutanoic acid, 5-hydroxymethyl-2-furoic acid, 2-keto-D-gluconic acid, threitol, 3,4,5-trihydroxy-pentanoic acid, allonic acid, D-glycero-L-manno-heptonic acid, idonic acid, and scopolin, but had negative correlations with glycine, histidine, arginine, alanine, methionine, and lysine. Malic acid showed strong positive correlations with citric acid, 2-hydroxybutanoic acid, 5-hydroxymethyl-2-furoic acid, 2-keto-D-gluconic acid, threitol, 3,4,5-trihydroxy-pentanoic acid, idonic acid, and scopolin.

Sucrose showed strong positive correlations with ribose, 2-keto-D-gluconic acid, D-glycero-L-manno-heptonic acid, fructose, glucose, galactose, arabinofuranose, glucopyranoside, D-glycero-L-manno-heptonic acid, mannonic acid, scopolin, and palatinose. Glycerol 3-phosphate and glycerol were placed close to each other in the metabolic network, because glycerol 3-phosphate is produced from glycerol by the enzyme glycerol kinase and the conversion of glycerol 3-phosphate to glycerol is catalysed by glycerol 3-phosphatase. Obviously, these data suggested that during fruit maturation, carbohydrate biosynthesis increased only when free amino acid accumulation decreased.

Discussion

Several promising functional genomic approaches, including transcriptomics, proteomics, and metabolomics, are employed in the systematic and comprehensive understanding of the complex events of life. Currently, whole genome information of strawberry, an octoploid species, is unavailable, restricting transcriptomic and proteomic studies. Metabolomics can be independent of genomic data, identifying sample constituents and the actual biochemical status of the tissues.

Despite the wide application of metabolomics in plant research, there are few published reports on the metabolite profiles of strawberry fruits. These reports either described non-targeted metabolomic analysis of dissected strawberry floral organs by the use of UPLC-qTOF-MS (Hanhineva *et al.*, 2008) or were focused on the metabolite composition of a fruit polar extract by use of Fourier transform ion cyclotron mass spectrometry (Aharoni *et al.*, 2002) and GC-MS (Aprea *et al.*, 2009). Thus, an extensive and comprehensive metabolite profile in strawberry fruits (including non-polar and polar phases) has not been described previously. This study provides a comprehensive, comparative analysis of the metabolite composition of strawberry fruits from seven different developmental stages. The resulting information is important for investigating the changes of key nutritional metabolites and for illustrating the mechanism of fruit quality formation.

A derivatization step is often necessary before GC-MS analysis in order to improve the volatility of polar metabolites. At the same time, due to the cyclic and open chain structures of sugars, silylation of monosaccharides without an oximation step leads to multiple peaks belonging to each individual sugar compound. By introducing an oximation step prior to silylation, cyclization is inhibited, resulting in fewer peaks per sugar. However, trimethylsilylation of some polar metabolites in this study still gave 2–3 peaks, such as in the case of fructose, galactose, and turanose (Figs 5, 7; Supplementary Table S1 at *JXB* online).

When data from polar extracts and non-polar extracts were combined, stage 1 (small green fruit) and stage 2 (large green fruit) have the most similar metabolite content of the seven developmental stages (Fig. 6). It is probable that both stages actually belong to the same fruit developmental phase (the cell division phase). After that, during 15–35 DPA, fruits were characterized by the accumulation of soluble sugars and organic acids, which contribute to the acquisition of the fleshy trait associated with cell expansion. Significant distinctions of metabolite composition between stage 4 (white fruit) and stage 5 (red-turning fruit) were detected based on metabolic clusters using PLS-DA (Fig. 6). This is due to the developmental transition from cell expansion to cell maturation (Knee *et al.*, 1977), in agreement with increases in pigment accumulation (Lopes *et al.*, 2007) and shifts in gene expression (Medina *et al.*, 1997; Manning, 1998; Nam *et al.*, 1999). Because many proteins serve as storage units of fruit tissues, in contrast to the accumulation of most other metabolites, the levels of free amino acids decreased sharply. In Fig. 8, the fruit samples from 25 DPA mapped with the samples from 30 DPA, reflecting the compositional similarities of amino acid levels between those two phases.

During the ripening phase (30–40 DPA), major pigments Pg3glc and Cy3glc, which change the colour of the flesh, increased drastically. In parallel, fructose, galactose, sucrose, citrate, and malate concentrations continue to increase, leading to a higher sugar/acid ratio, which is a major determinant of fruit taste. Additionally, there is a relationship between the softening of the strawberry fruit and the degradation of the middle lamella and cell wall, which mostly occurs during the last stages of the ripening process (Knee *et al.*, 1977). As a result of hydration of the cell walls (Knee *et al.*, 1977) and alterations in the cross-linking of carbohydrates (Huber, 1984; Manning, 1993), >70% of the polyuronide in the cell wall became freely soluble. The observed increases of galactose (1.5-fold) and arabinose (4.2-fold) in stage 6 are indicative of possible changes in cell wall structure linked with fruit softening. The increase in galactose during stage 6 followed by a decrease in stage 7 was also observed, which suggests that the process of fruit senescence was slower in the red-ripening stage than in the over-ripening stage.

The metabolite profiling analysis indicated the perturbation of several metabolic pathways, including ester biosynthesis, the shikimate pathway, the TCA cycle, and amino acid synthesis, during fruit growth and ripening. The amino acid alanine has been implicated in the formation of ethyl esters (Perez *et al.*, 1992), and two aromatic amino acids (phenylalanine and tyrosine) are known to be the precursors for the biosynthesis of anthocyanins and flavonoids through the shikimate pathway. Four central amino acids, glutamine, glutamate, aspartate, and asparagine (Gln, Glu, Asp, and Asn), are first derived from α-oxoglutarate and oxaloacetate in the TCA cycle (Fig. 9) and then converted into all other amino acids by various biochemical processes (Galili *et al.*, 2008). Taken together, the present data suggest that amino acid metabolism is central to fruit development and that the synthesis of storage proteins and other amino acid-derived compounds may have major impacts on the levels of free amino acids.

Volatile compounds play a dual role in fruit development and ripening, serving both as biological perfumes to entice living creatures, including humans, and as protectants against pathogens. Thus, the blends of flavour compounds produced by fruits constitute fruit organoleptic properties. The volatile components of strawberry aroma formed during ripening are the result of the combined perception of esters, alcohols, and alkanes. The formation of these compounds is closely correlated with the metabolic changes occurring during fruit maturation. The most likely precursors for the esters are lipids and amino acids. The degradation of fatty acids results in the production of volatile aldehydes, which are subsequently converted to alcohols and hexyl esters (Perez *et al.*, 1996). The last step in ester biosynthesis is catalysed by alcohol acyltransferases, which link alcohols to acyl moieties (Shalit *et al.*, 2001).

Alkanes also serve as volatile metabolites and contribute to fruit flavour. In ripening fruits of 'Kensington Pride' mango, the most abundant group of volatile compounds was alkanes, accounting for ~59% of the total identified compounds (Lalel *et al.*, 2003). The amount of most of the alkanes in strawberry increased from the small green fruit stage, reached the highest point at the white fruit stage, and gradually decreased through the red-ripening and over-ripening stages. Among these volatiles, it was found that 2,5-dimethyl-4-hydroxy-3(2H)-furanone is a major flavour compound because of its high concentration (up to 48.53 mg kg^{-1} strawberry fruit fresh weight) and its low odour threshold (10 ppb), in agreement with previous data (Larsen and Poll, 1992; Schwab and Roscher, 1997). Fructose-1,6-bisphosphate represents the major precursor of 2,5-dimethyl-4-hydroxy-3(2H)-furanone in strawberry fruits (Roscher *et al.*, 1998), while the enzyme quinone oxidoreductase was reported to be involved in the biosynthesis of 2,5-dimethyl-4-hydroxy-3(2H)-furanone (Raab *et al.*, 2006), which can be confirmed with the relatively high correlation between fructose and 2,5-dimethyl-4-hydroxy-3(2H)-furanone ($r=0.776$, $P=4.04\%$) in this study.

Compared with published strawberry metabolomics data, the present report covered more developmental stages and analysed more classes of metabolites encompassing all major types of small molecules. More importantly, the correlation analysis found many interesting connections between different metabolic pathways. For example, sucrose was found to be closely linked to the biosynthesis of other monosaccharides and disaccharides, which share the same Calvin cycle origin. Surprisingly, increased glycoside conjugation, such as 1-methyl-D-galactopyranoside, is negatively correlated with sucrose content, suggesting the carbon source at the mature fruit may be limited.

Several correlations in the metabolic network that are difficult to explain were also discovered. For example, fructose, galactose, and malic acid strongly correlated with 2-butyl-l-octanol accumulation. Gutierrez-Gonzalez *et al.* (2009) previously suggested that in seed, biosynthesis of sugar and starch might be in direct competition with the biosynthesis of fatty acid and lipid, which is the precursor of 2-butyl-l-octanol. Therefore, a reverse correlation between monosaccharides and long-chain fatty acid or alkanes is expected. This and other unexpected connections deserve further investigations.

In conclusion, by metabolic profiling of both polar and non-polar compounds, it was possible to monitor the alterations in several major groups of compounds during strawberry fruit growth and maturation. The discoveries in these pathways are consistent with strawberry pigment and flavour formation, using mostly materials from primary metabolism. Each stage of fruit development has its own unique metabolic profiles, with the most drastic changes occurring at the transition toward the red-ripened stage. Amino acid biosynthesis plays an important role in generating several classes of compounds related to the quality of the fruits. This information will help strawberry breeders detect and monitor key components that are important for these output traits.

Supplementary data

Supplementary data are available at *JXB* online.

Figure S1. Metabolite identification of fructose by GC-MS as an example. (A) MS spectrum of the targeted peak (RT=36.425 min) in the polar phase. The mass spectrum contains abundant ions at *m/z* 73.0 and *m/z* 103.1, which coincide with the two moieties (CH$_3$)$_3$Si and (CH$_3$)$_3$SiNO in trimethylsilyl-fructose. (B) MS spectrum of the fructose standard. (C) Proposed pathway of fructose trimethylsilylation. Fructose was derivatized through oximation and trimethylsilylation with the addition of methoxamine hydrochloride and BSTFA containing 1% TMCS, respectively.

Figure S2. Comparison of selected MS spectra from the targeted peaks in the polar extract and their authentic reference compounds. (A) MS spectrum of the targeted peak (RT=39.881 min) in the polar phase; (B) MS spectrum of the galactose standard; (C) MS spectrum of the targeted peak (RT=72.267 min) in the polar phase; (D) MS spectrum of the sucrose standard; (E) MS spectrum of the targeted peak (RT=30.407 min) in the polar phase; (F) MS spectrum of the citric acid standard; (G) MS spectrum of the targeted peak (RT=17.801 min) in the polar phase; (H) MS spectrum of the malic acid standard.

Figure S3. PCA scatter plot explaining the separation of polar phase samples from different strawberry fruit stages. The PCA vector plot using principle component 1 (P1) and principle component 2 (P2), representing 40.91% and 4.58%, of the variance between the data points, respectively. Each data point represents a linear combination of all the metabolites from a single extract. Each data point represents a unique compound, and several VIP compounds in Table 2 are labelled in this figure.

Figure S4. HPLC chromatograms showing the separation of different amino acid standards (A) and different amino acid components in a representative sample of strawberry fruit (B). Labelled peaks were identified as: 1, aspartic acid (Asp); 2, serine (Ser); 3, glutamic acid (Glu); 4, glycine (Gly); 5, histidine (His); 6, α-aminobutyric acid (NH$_3$); 7, arginine (Arg); 8, threonine (Thr); 9, alanine (Ala); 10, proline (Pro); 11, cysteine (Cys); 12, tyrosine (Tyr); 13, valine (Val); 14, methionine (Met); 15, lysine (Lys); 16, isoleucine (Ile); 17, leucine (Leu); 18, phenylalanine (Phe). The cysteine content is 10 μmol l^{-1} and the content of the other amino acid standards is 20 μmol l^{-1}.

Figure S5. Metabolite network analysis on strawberry fruit polar extracts. Several metabolites are labelled, including fructose, galactose, sucrose, citric acid, malic acid, and *myo*-inositol-1-phosphate. Metabolite network analysis clearly demonstrated that the four main components in the polar extracts (fructose, galactose, citric acid, and malic acid) had many linkages with other metabolites and could be placed in the core of the metabolic network.

Table S1. Tentatively identified metabolites in the untargeted GC-MS and targeted HPLC analysis and their abundance in different developmental stages of strawberry fruits.

Table S2. Metabolite correlation analysis of hexadecane, isobutyl phthalate, and 2-ethylhexyl phthalate from the cyclohexane extract.

References

Aaby K, Ekeberg D, Skrede G. 2007. Characterization of phenolic compounds in strawberry (*Fragaria × ananassa*) fruits by different HPLC detectors and contribution of individual compounds to total antioxidant capacity. *Journal of Agricultural and Food Chemistry* **55,** 4395–4406.

Aharoni A, Keizer LCP, Bouwmeester HJ, et al. 2000. Identification of the SAAT gene involved in strawberry flavor biogenesis by use of DNA microarrays. *The Plant Cell* **12,** 647–661.

Aharoni A, de Vos CH, Maliepaard CA, Kruppa G, Bino RJ, Goodenough D. 2002. Nontargeted metabolome analysis by use of fourier transform ion cyclotron mass spectrometry. *OMICS* **6,** 217–234.

Anttonen MJ, Hoppula KI, Nestby R, Verheul MJ, Karjalainen RO. 2006. Influence of fertilization, mulch color, early forcing, fruit order, planting date, shading, growing environment, and genotype on the contents of selected phenolics in strawberry (*Fragaria × ananassa* Duch.) fruits. *Journal of Agricultural and Food Chemistry* **54,** 2614–2620.

Aprea E, Biasioli F, Carlin S, Endrizzi I, Gasperi F. 2009. Investigation of volatile compounds in two raspberry cultivars by two headspace techniques: solid-phase microextraction/gas chromatography-mass spectrometry (SPME/GC-MS) and proton-transfer reaction-mass spectrometry (PTR-MS). *Journal of Agricultural and Food Chemistry* **57,** 4011–4018.

Atkinson CJ, Dodds PA, Ford YY, Le Mière J, Taylor JM, Blake PS, Paul N. 2006. Effects of cultivar, fruit number and reflected photosynthetically active radiation on *Fragaria× ananassa* productivity and fruit ellagic acid and ascorbic acid concentrations. *Annals of Botany* **97,** 429–444.

Blount JW, Masoud S, Sumner LW, Huhman D, Dixon RA. 2002. Over-expression of cinnamate 4-hydroxylase leads to increased accumulation of acetosyringone in elicited tobacco cell-suspension cultures. *Planta* **214,** 902–910.

Broeckling CD, Huhman DV, Farag MA, Smith JT, May GD, Mendes P, Dixon RA, Sumner LW. 2005. Metabolic profiling of *Medicago truncatula* cell cultures reveals the effects of biotic and abiotic elicitors on metabolism. *Journal of Experimental Botany* **56,** 323–326.

Chen F, Kota P, Blount JW, Sumner LW, Dixon RA. 2003. Profiling phenolic metabolites in transgenic alfalfa modified in lignin biosynthesis. *Phytochemistry* **64,** 1013–1021.

Fait A, Hanhineva K, Beleggia R, Dai N, Rogachev I, Nikiforova VJ, Fernie AR, Aharoni A. 2008. Reconfiguration of the achene and receptacle metabolic networks during strawberry fruit development. *Plant Physiology* **148,** 730–750.

Fiehn O. 2003. Metabolic networks of *Cucurbita maxima* phloem. *Phytochemistry* **62,** 875–886.

Galili S, Amir R, Galili G. 2008. Genetic engineering of amino acid metabolism in plants. *Advances in Plant Biochemistry and Molecular Biology* **1,** 49–80.

Gu L, Kelm MA, Hammerstone JF, Beecher G, Holden J, Haytowitz D, Prior RL. 2003. Screening of foods containing proanthocyanidins and their structural characterization using LC–MS/MS and thiolytic degradation. *Journal of Agricultural and Food Chemistry* **51,** 7513–7521.

Guo W, Sakata K, Watanabe N, Nakajima R, Yagi A, Ina K, Luo S. 1993. Geranyl 6-*O*-β-D-xylopyranosyl-β-D-glucopyranoside isolated as an aroma precursor from tea leaves for oolong tea. *Phytochemistry* **33,** 1373–1375.

Gutierrez-Gonzalez JJ, Wu X, Zhang J, Lee JD, Ellersieck M, Shannon GJ, Yu O, Nguyen HT, Sleper DA. 2009. Genetic control of soybean seed isoflavone content: importance of statistical model and epistasis in complex traits. *Theoretical and Applied Genetics* **119,** 1069–1083.

Hamzehzarghani H, Kushalappa AC, Dion Y, Rioux S, Comeau A, Yaylayan V, Marshall WD, Mather DE. 2005. Metabolic profiling and factor analysis to discriminate quantitative resistance in wheat cultivars against fusarium head blight. *Physiological and Molecular Plant Pathology* **66,** 119–133.

Hancock JF. 1999. *Strawberries*. Wallingford, UK: CABI Publishing.

Hanhineva K, Rogachev I, Kokko H, Mintz-Oron S, Venger I, Kärenlampi S, Aharoni A. 2008. Non-targeted analysis of spatial metabolite composition in strawberry (*Fragaria × ananassa*) flowers. *Phytochemistry* **69,** 2463–2481.

Heinonen IM, Meyer AS, Frankel EN. 1998. Antioxidant activity of berry phenolics on human low-density lipoprotein and liposome oxidation. *Journal of Agricultural and Food Chemistry* **46,** 4107–4112.

Honkanen E, Hirvi T. 1990. The flavour of berries. In: Morton ID, MacLeod AJ, eds. *Food flavours*. Amsterdam: Elsevier Scientific Publications, 125–193.

Huber DJ. 1984. Strawberry fruit softening: the potential roles of polyuronides and hemicelluloses. *Journal of Food Science* **49,** 1310–1315.

Hukkanen AT, Kokko HI, Buchala AJ, McDougall GJ, Stewart D, Kärenlampi SO, Karjalainen RO. 2007. Benzothiadiazole induces the accumulation of phenolics and improves resistance to powdery mildew in strawberries. *Journal of Agricultural and Food Chemistry* **55,** 1862–1870.

Kim JK, Bamba T, Harada K, Fukusaki E, Kobayashi A. 2007. Time-course metabolic processing in *Arabidopsis thaliana* cell cultures

after salt stress treatment. *Journal of Experimental Botany* **58,** 415–424.

Knee M, Sargent JA, Osborne DJ. 1977. Cell-wall metabolism in developing strawberry fruits. *Journal of Experimental Botany* **28,** 377–396.

Lalel HJD, Singh Z, Tan SC. 2003. Aroma volatiles production during fruit ripening of 'Kensington Pride' mango. *Postharvest Biology and Technology* **27,** 323–336.

Larsen M, Poll L. 1992. Odor thresholds of some important aroma compounds in strawberries. *Zeitschrift Lebensmittel-Untersuchung und -Forschung* **195,** 120–123.

Latrasse A. 1991. Fruits III. In: Maarse H, ed. *Volatile compounds in foods and beverages*. New York: Dekker, 329–387.

Lopes da Silva F, Escribano-Bailón MT, Pérez Alonso JJ, Rivas-Gonzalo JC, Santos-Buelga C. 2007. Anthocyanin pigments in strawberry. *LWT–Food Science and Technology* **40,** 374–382.

Maarse H. 1991. *Volatile compounds in foods and beverages*. New York: Marcel Dekker, 483–546.

Määttä-Riihinen KR, Kamal-Eldin A, Torronen AR. 2004. Identification and quantification of phenolic compounds in berries of Fragaria and Rubus species (family Rosaceae). *Journal of Agricultural and Food Chemistry* **52,** 6178–6187.

Manning K. 1998. Isolation of a set of ripening-related genes from strawberry: their identification and possible relationship to fruit quality traits. *Planta* **205,** 622–631.

Manning K. 1993. Soft fruits. In: Taylor JE, Tucker GA, eds. *Biochemistry of fruit ripening*. Cambridge, UK: Chapman and Hall, 347–373.

Mateo JJ, Gentilini N, Huerta T, Jimenez M, Distefano R. 1997. Fractionation of glycoside precursors of aroma in grapes and wine. *Journal of Chromatography A* **778,** 219–224.

Medina Escobar N, Cardenas J, Valpuesta V, Munoz Blanco J, Caballero JL. 1997. Cloning and characterization of cDNAs from genes differentially expressed during the strawberry fruit ripening process by a MAST-PCR-SBDS method. *Analytical Biochemistry* **248,** 288–296.

Moco SIA, Bino RJ, Vos de CH, Vervoort JJM. 2007. Metabolomics technologies and metabolite identification. *Trends in Analytical Chemistry* **26,** 855–866.

Nam YW, Tichit L, Leperlier M, Cuerq B, Marty I, Lelievre JM. 1999. Isolation and characterization of mRNAs differentially expressed during ripening of wild strawberry (*Fragaria vesca* L.) fruits. *Plant Molecular Biology* **39,** 629–636.

Perez AG, Rios JJ, Sanz C, Olias JM. 1992. Aroma components and free amino acids in strawberry variety Chandler during ripening. *Journal of Agricultural and Food Chemistry* **40,** 2232–2235.

Perez AG, Sanz C, Olias R, Rios JJ, Olías JM. 1996. Evolution of strawberry alcohol acyltransferase activity during fruit development and storage. *Journal of Agricultural and Food Chemistry* **44,** 3286–3290.

Perkins-Veazie P. 1995. Growth and ripening of strawberry fruit. *Horticultural Reviews* **17,** 267–297.

Raab T, Lopez-raez JA, Klein D, Caballero JL, Moyano E, Schwab W, Munoz-Blanco J. 2006. FaQR, required for the

biosynthesis of the strawberry flavor compound 4-hydroxy-2,5-dimethyl-3(2H)-furanone, encodes an enone oxidoreductase. *The Plant Cell* **18,** 1023–1037.

Raamsdonk LH, Teusink B, Broadhurst D, et al. 2001. A functional genomics strategy that uses metabolome data to reveal the phenotype of silent mutations. *Nature Biotechnology* **19,** 45–50.

Roessner-Tunali U, Hegemann B, Lytovchenko A, Carrari F, Bruedigam C. 2003. Metabolic profiling of transgenic tomato plants overexpressing hexokinase reveals that the influence of hexose phosphorylation diminishes during fruit development. *Plant Physiology* **133,** 84–99.

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

Roessner U, Wagner C, Kopka J, Trethewey RN, Willmitzer L. 2000. Simultaneous analysis of metabolites in potato tuber by gas chromatography–mass spectrometry. *The Plant Journal* **23,** 131–142.

Roscher R, Bringmann G, Schreier P, Schwab W. 1998. Radiotracer studies on the formation of 2, 5-dimethyl-4-hydroxy-3[2H]-furanone in detached ripening strawberry fruits. *Journal of Agricultural and Food Chemistry* **46,** 1488–1493.

Santos-Buelga C, Scalbert A. 2000. Proanthocyanidins and tannin-like compounds: nature, occurrence, dietary intake, and effects on nutrition and health. *Journal of the Science of Food and Agriculture* **80,** 1094–1117.

Sato S, Soga T, Tomita M. 2004. Simultaneous determination of the main metabolites in rice leaves using capillary electrophoresis mass spectrometry and capillary electrophoresis diode array detection. *The Plant Journal* **40,** 151–163.

Schauer N, Fernie AR. 2006. Plant metabolomics: towards biological function and mechanism. *Trends in Plant Science* **11,** 508–516.

Schauer N, Zamir D, Fernie AR. 2005. Metabolic profiling of leaves and fruit of wild species tomato: a survey of the *Solanum lycopersicum* complex. *Journal of Experimental Botany* **56,** 297–307.

Schwab W, Roscher R. 1997. 4-Hydroxy-3(2H)-furanones: natural and maillard products. *Recent Research Developments in Phytochemistry* **1,** 643–673.

Shalit M, Katzir N, Tadmor Y, Larkov O, Burger Y, Schalechet F, Lastochkin E, Ravid U, Amar O, Edelstein M. 2001. Acetyl-CoA:alcohol acetyl transferase activity and aroma formation in ripening melon fruits. *Journal of Agricultural and Food Chemistry* **49,** 794–799.

Steuer R, Kurths J, Fiehn O, Weckwerth W. 2003*a*. Interpreting correlations in metabolomic networks. *Biochemical Society Transactions* **31,** 1476–1478.

Steuer R, Kurths J, Fiehn O, Weckwerth W. 2003*b*. Observing and interpreting correlations in metabolomic networks. *Bioinformantics* **19,** 1019–1026.

Tagashira N, Plader W, Filipecki M, et al. 2005. The metabolic profiles of transgenic cucumber lines vary with different chromosomal locations of the transgene. *Cellular and Molecular Biology Letters* **10,** 697–710.

Terry LA, Chope GA, Bordonaba JG. 2007. Effect of water deficit irrigation and inoculation with *Botrytis cinerea* on strawberry (*Fragaria*

× *ananassa*) fruit quality. *Journal of Agricultural and Food Chemistry* **55,** 10812–10819.

Wang SY, Jiao H. 2000. Scavenging capacity of berry crops on superoxide radicals, hydrogen peroxide, hydroxyl radicals and singlet oxygen. *Journal of Agricultural and Food Chemistry* **48,** 5677–5684.

Wang SY, Lin HS. 2000. Antioxidant activity in fruits and leaves of blackberry, raspberry and strawberry varies with cultivar and developmental stage. *Journal of Agricultural and Food Chemistry* **48,** 140–146.

Weckwerth W, Loureiro M, Wenzel K, Fiehn O. 2004. Differential metabolic networks unravel the effects of silent plant phenotypes. *Proceedings of the National Academy of Sciences, USA* **101,** 7809–7814.

Zabetakis I, Holden MA. 1997. Strawberry flavor: analysis and biosynthesis. *Journal of the Science of Food and Agriculture* **74,** 421–434.