Journal of Experimental Botany, Vol. 60, No. 7, pp. 2139–2154, 2009
doi:10.1093/jxb/erp086 Advance Access publication 3 April, 2009
This paper is available online free of all access charges (see http://jxb.oxfordjournals.org/open_access.html for further details)

Journal of Experimental Botany
www.jxb.oxfordjournals.org

RESEARCH PAPER

Metabolic characterization of loci affecting sensory attributes in tomato allows an assessment of the influence of the levels of primary metabolites and volatile organic contents

Maria Inés Zanor[1,*], José-Luis Rambla[2], Jamila Chaïb[3], Agnes Steppa[1], Aurora Medina[2], Antonio Granell[2], Alisdair R. Fernie[1,†] and Mathilde Causse[3]

[1] Max-Planck-Institute for Molecular Plant Physiology, Am Mühlenberg 1, D-14476 Potsdam-Golm, Germany
[2] Universidad Politecnica Valencia, CSIC, Instituto de Biología Molecular y Celular Plantas (IBMCP), Avda de los Naranjos s/n, Valencia 46022, Spain
[3] INRA, UR1052, Unité de Génétique et Amélioration des Fruits et Légumes, Domaine Saint-Maurice, BP94, F-84143 Montfavet Cedex, France

Received 8 February 2009; Revised 2 March 2009; Accepted 2 March 2009

Abstract

Numerous studies have revealed the extent of genetic and phenotypic variation between both species and cultivars of tomato. Using a series of tomato lines resulting from crosses between a cherry tomato and three independent large fruit cultivar (Levovil, VilB, and VilD), extensive profiling of both central primary metabolism and volatile organic components of the fruit was performed. In this study, it was possible to define a number of quantitative trait loci (QTLs) which determined the levels of primary metabolites and/or volatile organic components and to evaluate their co-location with previously defined organoleptic QTLs. Correlation analyses between either the primary metabolites or the volatile organic compounds and organoleptic properties revealed a number of interesting associations, including pharmaceutical aroma–guaiacol and sourness–alanine, across the data set. Considerable correlation within the levels of primary metabolites or volatile organic compounds, respectively, were also observed. However, there was relatively little association between the levels of primary metabolites and volatile organic compounds, implying that they are not tightly linked to one another. A notable exception to this was the strong association between the levels of sucrose and those of a number of volatile organic compounds. The combined data presented here are thus discussed both with respect to those obtained recently from wide interspecific crosses of tomato and within the framework of current understanding of the chemical basis of fruit taste.

Key words: Metabolite profiling, QTL sensory profiling, Tomato, Volatile profiling.

Introduction

Human perception of flavour involves the integration of multiple signals emanating from taste and olfactory receptors. In tomato, as in most fruits, flavour is largely dependent on sugar and acid contents, but also on the sugar/acid ratio (Dennison *et al.*, 1953; Stevens 1972; Saliba-Colombani *et al.*, 2001). However, whilst taste receptors clearly respond to relatively few cues, olfactory receptors respond to thousands of chemicals and as such are thought to be responsible for the vast diversity of unique food flavours (Goff and Klee, 2006; Tieman *et al.*, 2006*a*). In the case of tomato fruits, ~400 volatile organic compounds have been identified (Petro-Turza, 1987), between 15 and 20 of which are thought to constitute the flavour of fresh tomatoes (Buttery *et al.*, 1971; Baldwin

* Present address: Instituto de Biología Molecular y Celular de Rosario (IBR) Suipacha 531, S2002LRK Rosario, Argentina.
† To whom correspondence should be addressed. E-mail: fernie@mpimp-golm.mpg.de

et al., 2000). These volatile compounds are generally derived from various precursors including fatty acids, carotenoids, and amino acids. However, the exact definition of the biosynthetic pathways of many of them remains elusive (Tieman *et al.*, 2006*a*). In addition to the chemical components of fruit quality, physical components related to texture are of crucial importance to the consumer (Causse *et al.*, 2003; Serrano-Megias and Lopez-Nicolas, 2006; Chaïb *et al*, 2007). Fruit texture is composed of many traits including flesh firmness, mealiness, meltiness, juiceness, and crispness (Harker *et al.*, 1997; Redgwell and Fischer, 2002; Szczesniak, 2002). During fruit ripening, major changes in texture occur. Fruit softening has a major impact on many aspects of post-harvest physiology, including transport, shelf life, and disease resistance (Brummell and Harpster, 2001; Saladie *et al.*, 2006).

Given that consumers have complained about tomato flavour for >10 years in Europe (Decoene, 1995; Janse and Schols, 1995), the USA (De Giglio, 2003), and Australia (Ratanachinakorn *et al.*,1997), much research attention has focused on ways to improve it. As a first step in this process a number of surveys of natural variation in the chemical composition of tomatoes have been carried out either on the cultivar/species basis (Schauer *et al.*, 2005*b*; Spencer *et al.*, 2005; Tikunov *et al.*, 2005; Fernie *et al.*, 2006), or utilizing either recombinant inbred or introgression lines (Chaib *et al.*, 2006, 2007; Schauer *et al.*, 2006, 2008; Tieman *et al.*, 2006*b*; Hovav *et al.*, 2007). Several of these studies have identified genomic loci controlling the levels either of sugars and organic acids or of volatiles (Saliba-Colombani *et al.*, 2001; Causse *et al.*, 2002; Tieman *et al.*, 2006*b*; Schauer *et al.*, 2006, 2008), whilst other studies have concentrated on more physical aspects of organoleptic quality (Lecomte *et al.*, 2004; Chaïb *et al*, 2007). In the current study, the metabolite composition of quantitative trait loci near isogenic lines (QTL-NILs) that had previously been demonstrated, by use of a trained tasting panel, to possess characteristic organoleptic properties (Chaïb *et al*, 2006) were evaluated. For this purpose, both polar primary metabolites and volatile organic compounds in the lines were evaluated using well-established GC-MS-based profiling methods for each type of compound. In total, the levels of ~100 metabolites were determined and it was possible to evaluate co-localization and correlation of changes in these metabolic traits with changes in the previously determined organoleptic traits. Data are discussed with respect to current models of determinants of fruit organoleptic quality and its underlying molecular basis.

Materials and methods

Plant material

The experiments were performed on parental lines and two types of introgressed lines in different genetic backgrounds: genotypes combining five regions of interest for fruit quality and QTL-NILs carrying one introgressed region of chromosome 1, 2, 4, and 9 (two regions 9A and 9B). The five regions carried several QTLs involved in fruit quality (see Fig. 3, Causse *et al.*, 2002). The initial QTL analysis was performed on a population of recombinant inbred lines (RILs) developed from an intraspecific cross between Cervil (a cherry tomato, *Solanum lycopersicum*, var. *cerasiforme*) with 7 g fruits, a good taste, and a high aroma intensity, and Levovil (a *S. lycopersicum* line) with 125 g fruits and a common taste (Causse *et al.*, 2002). Based on the QTL map, five regions (located on chromosomes 1, 2, 4, and 9, respectively) were introgressed in the Levovil genetic background. A QTL for titratable acidity was detected in region 1, QTLs for sweetness, tomato aroma intensity, mealiness, and meltiness were detected in region 2, a QTL for mealiness and several QTLs for volatiles were detected in region 4, QTLs for sourness, tomato aroma intensity, mealiness, meltiness, and flesh firmness were detected in region 9A, and a QTL for pharmaceutical aroma was detected in region 9B. QTLs for physical and chemical traits were also detected in these regions. The introgressed lines were produced as described in Chaib *et al.* (2006). Briefly, as the favourable alleles for fruit quality were conferred by the C parent in most of the cases, the cherry tomato alleles at the five regions were introgressed into large fruit genotypes in order to obtain QTL-NILs. A single RIL with C alleles at the five regions was used as the donor parent of the breeding programme. The same marker-assisted backcross programme was performed with three different recipient lines, kindly provided by Vilmorin: Levovil, VilB, and VilD, hereafter L, B and D, respectively. As the donor parent contained 47% of recipient genome L, the first cross with each recipient line was considered as a BC1. The BC1 progeny was genetically homogenous; it was thus backcrossed without any selection to the recipient line to produce a BC2 population. Almost 300 plants were grown for each background, and, after a marker-assisted selection step, one BC2 individual was selected and backcrossed again to produce a BC3 population. Similarly, one BC3 individual was selected and three selfing generations were performed. In each BC3S1 population, the segregation of markers in the five regions of interest was comparable with that of an F_2 population. Then, BC3S3 lines with homozygous alleles at the five regions were selected and BC3S3 lines carrying C alleles at a single introgressed region were evaluated. These lines were nearly isogenic to their recipient line and were thus called QTL-NILs (Van Berloo *et al.*, 2001). The QTL-NILs were named with a letter corresponding to their genetic background and a number for the QTL region carried. For example, the line carrying the C allele at the region of interest on chromosome 2 with a genetic background L was denoted L2. In each genetic background, a line was obtained for each QTL region, with the exception of NIL-B9A that contained a C fragment introgressed on chromosome 1. The lines combining the five regions in the Levovil and VilB genetic background were named Lx and Bx, respectively.

Plant growth conditions trials

Three trials were performed during spring 2004, 2005, and 2006 in a heated glasshouse in Avignon (France, 43°55′N; 4°52′E). Planting took place on February at a density of 3.2 plants m^{-2}, and the day–night temperature set-point was 24–16 °C. Plant nutrition and chemical pest and disease control followed commercial practices and plants were grown on a single vine. From anthesis of the first truss, flowers were pollinated with an electrical shaker every 2–3 d. In each trial, the parental lines, the lines combining the five regions, and the QTL-NILs in the three genetic backgrounds were grown. Each line was represented by six plants grown in a fully randomized design. Several types of analyses were performed on red ripe tomatoes: physical measurements, sensory profiling, metabolic profiling, and volatile profiling.

Physical and physiological measurements

Red ripe fruits were harvested on the six plants of each line twice a week for 6 weeks. For metabolic profiles, six fruits per line were peeled and pericarp maintained frozen at –80 °C. For volatiles another six fruits per line were used and sections of the fruit were stored at –80 °C until further use.

Sensory profiling

Sensory profiles were obtained in 2004. Red ripe tomatoes were harvested in the morning of the day of the tasting, and homogeneous fruit samples were selected and stored at 20 °C in an air-conditioned room. The sensory panel was composed of 15 judges, who had previously been trained in the quantitative description of tomato attributes according to selection trials based on French norms (ISO8586-1, AFNOR V09-003). For each line, fruits were tasted twice by each judge, giving 30 scores per genotype. Fifteen sessions took place in a sensory analysis laboratory (AFNOR norm V09-105), on 2 d per week, and eight fruits were tasted by each judge on each occasion. The attributes chosen were colour intensity and heterogeneity, ribbed and translucent fruit intensity, to describe aspect, typical odour, sourness and sweetness, metal aroma, global aroma intensity, typical tomato aroma, pharmaceutical aroma, and firmness, juiciness, fleshiness, mealiness, and embarrassing skin to describe fruit texture. Each descriptor was scored on a 10-point scale.

Primary metabolite analysis

The relative levels of metabolites were determined using the GC-MS protocol exactly as described in Lisec *et al.* (2006) with the exceptions that the method was optimized for tomato fruit (Schauer *et al.*, 2006) and the mass spectra were cross-referenced with those in the Golm Metabolome Database (Kopka *et al.*, 2005; Schauer *et al.*, 2005*a*). The absolute concentrations of several metabolites were determined by comparison with calibration standard curve response ratios of various concentrations of standard substance solutions, including the internal standard ribitol (Roessner-Tunali *et al.*, 2003).

Volatile analysis

Fruit volatile analysis was performed essentially as described in Tikunov *et al.* (2005), with minor variations. Frozen tomato samples were milled in liquid nitrogen. A 1 g aliquot of the frozen fruit powder was weighed in a 7 ml vial, and the vial was sealed, and incubated at 37 °C for 10 min. An EDTA-NaOH water solution was prepared by adjusting 100 mM EDTA to a pH of 7.5 with NaOH. Then 1 ml of the EDTA-NaOH solution was added to the sample to a final EDTA concentration of 50 mM. A 2.2 g aliquot of solid $CaCl_2{\cdot}2H_2O$ was then immediately added. The closed vials were agitated and sonicated for 5 min. A 1 ml aliquot of the pulp was transferred into a 22 ml crimp cap vial (Perkin-Elmer), capped, and used for HS-SPME-GC-MS analysis. The vials were tempered at 50 °C for 10 min. The volatiles were then extracted by exposing a 65 μm polydimethylsiloxane-divinylbenzene SPME fibre (Supelco) to the vial headspace for 20 min under continuous agitation and heating at 50 °C. The fibre was manually inserted into a Clarus 500 (Perkin-Elmer) injection port and volatiles were desorbed for 1 min at 250 °C. Chromatography was performed on a ZB-5 (30 m×0.25 mm×0.25 μm) column with helium as carrier gas, at a constant flow of 1.2 ml min^{-1}. The GC interface and MS source temperatures were 260 °C and 180 °C, respectively. The oven programming conditions were 40 °C for 2 min, 5 °C min^{-1} ramp until 180 °C, then a 15 °C min^{-1} ramp until 250 °C, and a final hold at 250 °C for 4 min. The total run time, including oven cooling, was ~60 min. Mass spectra in the 35–250 *m*/*z* range were recorded by a Clarus 500 electron impact MS (Perkin-Elmer) at a scanning speed of five scans s^{-1} and an ionization energy of 70 eV. The chromatography and spectral data were evaluated using TurboMass software version 5.0 (Perkin-Elmer).

Data analysis

Statistical analyses were performed using either R statistical software or Microsoft Excel 7.0 (Microsoft, 2000). If two observations are described as different this means that their difference was determined to be statistically significant ($P < 0.05$) by the performance of Student's *t*-tests. The QTLs were evaluated by using Student's *t*-tests at a significance threshold of 0.05 to compare statistically each trait of each introgression line with its respective reference control. Principal component analysis was performed by means of SIMCA-P 11 software (Umetrics). Pearson correlation coefficients were calculated using the embedded CORREL function in Microsoft Excel 7.0 (Microsoft, 2000).

Heat map

Heat maps were calculated using the 'heatmap' module of the statistical software environment R (http://www.r-project.org) version 1.9. False colour imaging was performed on the

log2-transformed data. Regions of red and blue indicate negative or positive correlation between traits as depicted in the reference colour bar.

Results

Elite tomato lines harbour clear metabolic differences

Given that both previous sensory profiling results (Saliba-Colombani *et al.*, 2001; Causse *et al.*, 2003; Lecomte *et al.*, 2004) and common perception suggest that the cherry tomatoes are tastier than the large-fruited tomatoes, it was decided to analyse the basis of these differences at the metabolic level. For this purpose, an established GC-MS-based metabolite profiling method (Fernie *et al.*, 2004; Lisec *et al.*, 2006) was applied to the four parental lines used in this study [the cherry tomato line Cervil (C) and the large-fruited lines Levovil (L), VilB (B), and VilD (D)]. This analysis revealed profound differences between the lines in the levels of several metabolites. The initial focus was on the major sugar and acid contents (Fig. 1A). As could be anticipated, there were huge differences in sugar and acid levels between the three elite lines and the cherry tomato line, with the latter displaying greater levels of the major soluble sugars (sucrose, glucose, and fructose) whilst the larger fruited tomatoes had higher levels of malate and lower levels of citrate. In line with this observation, the sugar/acid ratio of the parental lines (calculated as μmol gFW^{-1} of sucrose, glucose, and fructose versus μmol gFW^{-1} of citrate and malate) was highest in the cherry variety (8.5) and lowest in the L variety (L=0.9; B=2.4; D=3.2). A more detailed analysis of the metabolite profiles of the parental lines revealed that many other metabolites were present at significantly different levels between the lines. One-way analysis of variance (ANOVA) tests revealed additional significant differences in the abundance of maltose, trehalose, arabinose, xylose, rhamnose, ribose, isocitrate, citramalate, malate, α-ketoglutarate, proline, valine, alanine, β-alanine, glutamate, serine, threonine, and phenylalanine between the parental lines. These data are presented in Table 1A which shows the fold changes observed in the levels of primary metabolites between each of the large-fruited cultivatrs and the cherry cultivar. It is well known that aroma makes a major contribution to the human perception of flavour (Goff *et al.*, 2006); therefore, analysis of volatile organic compounds was also conducted on the lines C, L, and B. This analysis revealed huge differences between the cherry variety C and the large-fruited varieties L and B, including changes in the levels of volatiles thought to be relevant for the definition of tomato aroma. The most prominent differences were in 2-phenylethanol, which was present at 6- to 13-fold higher levels in the C variety, and for a group of phenolic derivatives: eugenol, methylsalicylate, ethylsalicylate, and guaiacol, found at levels 20- to 100-fold lower than those observed in the large-fruited lines (Fig. 1B and Table 1B). Many other volatiles showed statistically

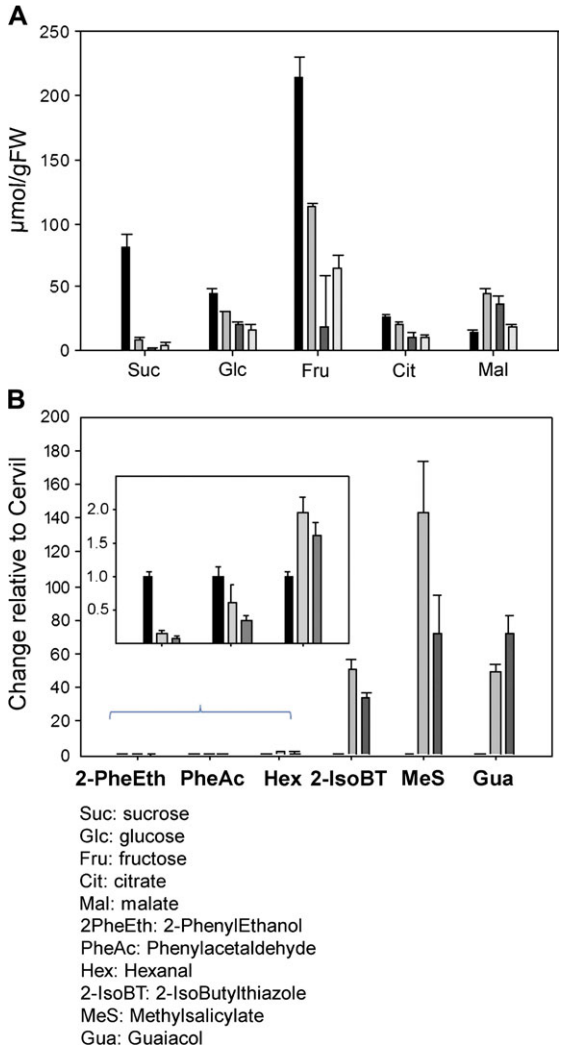

Fig. 1. Metabolic analysis of the parental lines. (A) Quantitative determination of the concentration of selected primary metabolites: sucrose, glucose, fructose, malate, and citrate in samples harvested in 2004. Cervil (black bars), VilB (light grey), Levovil (dark grey bars), and VilD (very light grey bars). Each bar represents the mean ±SE of six independent biological determinations. (B) Relative changes of 2-phenylethanol, phenylacetaldehyde, hexanal, 2-isobutylthiazole, methylsalicylate, and guaiacol in samples harvested in 2005. Cervil (black bars), VilB (light grey), and Levovil (dark grey). Data are normalized to the mean response for Cervil. Each bar represents the mean ±SE of five biologically independent replicates.

significant different levels between C and the other lines, such as terpineol, linalool, (E)-2-octenal, hexanal, (E)-2-pentenal, 1-penten-3-ol, 2-methylbutanol, (E)-2-methyl-2-butenal, 2-methylpropanal, benzaldehyde, phenylacetaldehyde, and 2-isobutylthiazole.

Analysis of metabolic variation in tomato lines pre-selected for their organoleptic properties

Having established that the elite lines displayed considerable metabolic variation, the primary metabolite content

Table 1A. Fold changes in the primary metabolites relative to Cervil in the parental lines

Data are normalized to the mean response calculated for the Cervil line. Values are presented as the mean of six biologically independent determinations. Those metabolites which were significantly different from the Cervil line (P <0.05) by the performance of Student's t-tests are marked in bold. nd indicates that metabolites were not detected.

	VilB	Levovil	VilD
Alanine	**0.17**	**0.25**	0.28
β-Alanine	**2.65**	0.33	1.83
Arginine/ornithine	0.86	**0.31**	0.71
Asparagine	**0.50**	**0.18**	**0.26**
Aspartate	1.52	**0.32**	**0.26**
γ-Aminobutyrate	**2.91**	0.84	1.34
Glutamate	**0.34**	**0.15**	**0.20**
Glutamine	0.40	0.46	**0.10**
Homoserine	1.01	0.47	0.64
Isoleucine	1.63	0.51	0.66
Lysine	**0.05**	0.68	**0.05**
Phenylalanine	0.89	**0.44**	**0.29**
Proline	**0.08**	**0.00**	**0.06**
Serine	**2.24**	1.68	**2.96**
Threonine	**2.12**	0.72	1.24
Valine	**3.06**	0.72	1.37
Arabinose	**0.58**	**0.58**	**0.34**
Fructose	**0.52**	**0.08**	**0.30**
Fructose-6-P	0.83	**0.37**	**0.44**
Gentiobiose	1.00	**0.06**	0.71
Glucose	**0.65**	**0.46**	**0.38**
Glucose-6-P	0.82	**0.02**	**0.39**
Maltose	1.09	**0.26**	**0.30**
Rhamnose	**0.61**	**0.71**	**0.37**
Ribose	**0.15**	**0.35**	**0.11**
Sucrose	**0.10**	**0.02**	**0.05**
Trehalose	**0.53**	**0.21**	**0.27**
Xylose	**0.69**	**2.54**	**0.38**
Glycerol	0.73	**0.12**	**0.47**
Myo-inositol	0.88	**0.33**	**0.35**
α-Ketoglutarate	**0.06**	**0.01**	**0.04**
Benzoate	0.87	0.58	**0.55**
Citramalate	**0.70**	**0.48**	**0.48**
Citrate	**0.73**	**0.39**	**0.35**
Dehydroescorbate	**0.33**	**0.38**	**0.20**
Fumarate	1.53	**1.76**	0.88
Galacturonate	**0.18**	1.14	**0.26**
Gluconate	**1.97**	1.15	1.13
Glycerate	**1.54**	**2.78**	**2.02**
Glycolate	0.91	**0.63**	**0.53**
Isocitrate	**0.37**	**0.19**	**0.29**
Malate	**3.26**	**2.63**	1.31
Nicotinate	**17.41**	nd	**7.47**
Phosphorate	0.86	**0.49**	**0.54**
Pyroglutamate	**0.44**	**0.28**	**0.28**
Quinate	**0.34**	**0.40**	**0.27**
Saccharate	**3.83**	**2.34**	**2.23**
Shikimate	0.56	2.78	0.06
Succinate	**0.46**	**0.40**	**0.31**
Threonate	**0.37**	**0.68**	**0.38**
FA 16:00	**2.11**	0.98	1.12
FA 18:00	**2.07**	0.97	1.14

Table 1B. Fold changes in the volatiles relative to Cervil in the parental lines

Data are normalized to the mean response calculated for the Cervil line. Values are presented as the mean of five biologically independent determinations. Those metabolites which were significantly different from the Cervil line (P <0.05) by the performance of Student's t-tests are marked in bold. Data for eugenol and ethylsalicylate should be considered as higher than the value present in the respective parental lines since these compounds were not present in the Cervil parental line.

	VilB	Levovil
2-Methyl-1-propanol	**0.31**	**0.27**
3-Methylbutanal	0.81	1.25
Butanol	0.42	0.82
1-Penten-3-ol	**0.65**	**0.60**
1-Penten-3-one	**0.69**	0.69
Pentanal	0.76	**0.55**
2-Ethylfuran	**0.60**	0.74
3-Methyl butanenitrile	0.98	**2.02**
3-Methylbutanol	**0.28**	0.58
2-Methyl-1-butanol	**0.40**	**0.31**
(E)-2-Methyl-2-butenal	**0.20**	**0.18**
(E)-2-Pentenal	**0.53**	**0.47**
1-Pentanol	**0.62**	**0.68**
(Z)-3-Hexenal	0.87	1.01
Hexanal	**1.97**	**1.59**
3-Methylbutanoic acid	**0.44**	0.52
(E)-2-Hexenal	**0.41**	**0.52**
(Z)-3-Hexen-1-ol	1.18	**2.06**
Pentanoic acid	**1.89**	1.05
(E,E)-2,4-Hexadien-1-al	0.83	1.01
α-Pinene	0.82	**0.33**
(E)-2-Heptenal	**0.72**	**0.64**
Benzaldehyde	**1.94**	**2.51**
6-Methyl-5-hepten-2-one	1.39	2.71
2-Pentylfuran	0.92	0.86
Hexanoic acid	1.79	0.76
Octanal	2.15	1.58
Benzylalcohol	**15.94**	**25.59**
2-Isobutylthiazole	**51.02**	**33.96**
Phenylacetaldehyde	0.63	**0.34**
(E)-2-Octenal	**0.42**	**0.55**
Acetophenone	2.47	2.54
p-Tolualdehyde	2.23	1.31
Guaiacol	**48.19**	**64.76**
Linalool	2.39	**9.91**
Nonanal	2.21	1.79
2-Phenylethanol	**0.17**	**0.10**
2-Ethyl-hexanoic acid	1.24	0.84
Benzylnitrile	0.87	0.87
Octanoic acid	1.64	1.08
Terpineol	**3.56**	**9.27**
Methyl salicylate	**142.88**	**69.15**
Geranial	1.95	**3.33**
Ethylsalicylate	>**29.55**	>**120.93**
1-Nitro-2-phenylethane	2.06	1.53
(E,E)-2,4-Decadienal	**0.24**	**0.41**
Eugenol	>**53.48**	>**45.62**
β-Damascenone	**1.63**	0.90
Geranylacetone	0.66	0.97
β-Ionone	0.63	1.39

of a subset of tomato lines resulting from their crossings which had been selected on the basis of their organoleptic properties (Lecomte *et al.*, 2004) were next evaluated. These lines consisted of marker-defined introgressions of five regions, controlling fruit quality variation, from the cherry tomato into each of the large-fruited lines. Lines in all three genetic backgrounds were evaluated in the first year but, due to the relatively low metabolic variation of the lines in the D background (see Supplementary Table S1 available at *JXB* online) subsequent studies were focused only on lines carrying the L and B backgrounds. The lack of phenotypic variation in the D background lines is largely in accordance with results of previous studies in suggesting unfavourable interaction on introgression of genome regions of C into the D variety (Lecomte *et al.*, 2004). A total of 45 primary metabolites were accurately quantified in every chromatogram. These compounds included most plant amino and organic acids, sugars, sugar alcohols, and fatty acids. The range of content of specific metabolites in the introgression lines was generally within that observed between the parental controls. In B background lines, only a relatively small number of metabolites exhibited transgressive behaviour in both harvests. These included glucose (which exhibited a range of relative levels of 0.54–1.49 in comparison with the recipient genotype control), aspartate (0.55–1.13), gluconate (0.00–1.33), β-alanine (0.79–4.06) and myo-inositol (0.52–1.07). All other metabolites only displayed transgressive behaviour either in a single harvest or not at all (see Table 2 for details). The occurrence of transgressive behaviour was even rarer in the L background and only reproducible in the case of alanine (which exhibited a range of relative levels of 0.58–9.58 in comparison with the recipient genotype control; Table 3).

Comparison of individual changes in primary metabolite content between the two harvests revealed that the data sets are generally in very high accordance, indicating that the observed changes are probably due to quantitative genetic factors. For subsequent analysis, the mean change between the two harvests was used since this allows a greater confidence that the changes reported are due to genetic rather than environmental factors. Whilst it is clearly difficult to display such a large data set in a truly quantitative manner, it can be stated that the mean difference in the content of any given metabolite ranged between 0.4 and 38.1 times the value observed in the L line for the L genotypes and between 0.3 and 9.7 times the value observed in the B line for the B genotypes. The metabolic changes observed in the hybrids, LxC and BxC, were similar in trend, but of more moderate magnitude, to the changes observed between the parental lines (Tables 2, 3). QTLs were determined by using Student's *t*-tests at a significance threshold of 0.05 in order to compare statistically every trait of each introgression line with its respective recipient genotype. Using this criterion, 35 single-trait metabolite QTLs were identified in the L background and 16 in the B background (see Fig. 2, although those for the introgression of chromosome 2 into the L background should be regarded as putative, since they only represent a single year analysis). Although most of the QTLs presented here were previously unknown, several, including those for sucrose and malate, have already been documented in the literature either in studies using the population described here or in studies reliant on the *S. pennellii* introgression line populations (Causse *et al.*, 2004; Schauer *et al.*, 2006, 2008). The number of QTLs was similar irrespective of the background into which the C genome segments were introgressed. Moreover, the F_1 hybrids between C and both L and B were largely equivalent with respect to the degree of metabolic changes observed [displaying changes in ~50% of traits (52% for L and 54% for B)].

The lines carrying the five introgressed segments simultaneously and hence the highest proportion of the parental cherry Cervil genome (Lx and Bx) showed a similar percentage of overall changes (~36% for Lx and 32% for Bx). Figure 2 shows the full list of QTLs (and, in the case of the Levovil introgression of chromosome 2, for which replicate data were not obtained, putative QTLs) for metabolite content, volatile content, and organoleptic properties analysed in the NILs. These QTLs were compared with the QTLs detected in a recombinant inbred population derived from the cross of Cervil and Levovil (Causse *et al.*, 2002). QTLs for sucrose were found in L1 and L2, which have previously been documented to display fruit sweetness QTLs. When the co-localization behaviour of the metabolites themselves is assessed, clustering of QTLs of metabolites of similar chemical structure is clearly visible, as would be expected both from previous studies of other traits in tomato (Causse *et al.*, 2002) and from studies of metabolic traits in both tomato and *Arabidopsis* (Schauer and Fernie, 2006; Lisec *et al.*, 2008; Rowe *et al.*, 2008).

Variation in volatile organic compound content in tomato lines pre-selected for their organoleptic properties

Having assessed the level of variation of primary metabolites in these lines, attention was next focused on the levels of volatile organic compounds. For this purpose, only L and B lines were studied. As for the primary metabolites, these compounds were measured in two different harvests—those of the 2005 and 2006 seasons (due to logistical difficulties it was not possible to perform these experiments in the exact same harvests; however, the close agreement of the primary metabolite results in the two harvests described above render this unproblematic). Fifty volatile organic compounds were accurately quantified by means of a HS-SPME-GC-MS method. In contrast to the observations for primary metabolites, many of the volatiles exhibited a transgressive behaviour. Guaiacol, (E)-2-pentenal, 1-pentanol, (Z)-3-hexenal, *p*-tolualdehyde, 3-methylbutanoic acid, and 2-pentylfuran showed transgressive behaviour in both genetic backgrounds analysed (Tables 4, 5). Additionally, 3-methylbutanal, 1-penten-3-one, 3-methylbutanenitrile, 3-methylbutanol, 2-methyl-1-butanol,

Table 2. Metabolic analysis of the lines derived from the cross between VilB and Cervil parents

Values are presented as the mean of six biologically independent determinations. The fold changes are relative to the VilB parent. In bold are those values which were significantly different with $P < 0.05$ by the performance of Student's *t*-tests.

	Harvest 2004						Harvest 2005					
	B1	**B2**	**B4**	**B9a**	**B9b**	**CxB**	**B1**	**B2**	**B4**	**B9a**	**B9b**	**Bx**
β-Alanine	0.95	1.04	0.68	**1.64**	1.18	**0.51**	1.34	**4.06**	0.79	1.34	1.39	**2.21**
Alanine	2.96	**2.37**	0.79	**2.32**	1.59	**2.41**	1.24	1.16	1.09	**2.17**	**1.84**	0.69
Asparagine	0.83	1.51	0.71	0.97	1.10	1.46	3.19	1.63	1.16	0.73	0.92	1.27
Aspartate	0.76	0.80	0.85	1.03	0.70	0.61	0.87	0.55	0.88	1.13	0.94	0.74
Cysteine	0.94	1.33	0.78	1.57	nd	0.29	0.77	0.90	0.77	0.69	0.61	0.58
γ-Aminobutyrate	0.88	0.92	0.92	1.15	0.79	**0.61**	**0.61**	0.91	**0.57**	**0.61**	0.69	**0.44**
Glutamate	1.34	0.97	0.92	1.08	0.67	1.51	1.04	1.00	0.96	1.12	1.20	1.08
Glutamine	0.92	1.09	0.67	1.04	0.95	nd	1.09	1.02	0.95	0.95	1.01	0.95
Glycine	0.78	1.22	0.66	1.20	1.49	0.23	0.82	1.56	0.91	1.64	2.24	0.65
Homoserine	0.66	0.77	0.53	0.66	0.65	0.57	0.40	0.35	0.87	0.41	0.35	0.39
Isoleucine	**0.55**	0.92	0.73	0.70	0.79	**0.59**	**0.76**	1.18	0.96	0.94	1.14	0.76
Lysine	0.85	0.89	1.00	0.71	1.78	27.82	0.70	0.53	1.28	0.44	**0.50**	0.65
Phenylalanine	0.79	1.16	0.61	0.72	0.73	**0.60**	0.88	1.13	0.94	0.92	1.24	1.02
Proline	**3.86**	**1.80**	0.63	**3.61**	1.10	**4.46**	**5.64**	**1.69**	1.07	**2.52**	**1.72**	**7.54**
Putrescine	**1.69**	1.15	1.83	**2.25**	1.57	**2.56**	1.15	1.06	**1.60**	**1.64**	1.19	1.75
Pyroglutamate	1.13	1.00	0.92	1.04	0.88	0.55	1.04	0.97	1.03	1.04	1.03	1.03
Serine	0.77	1.07	0.90	0.99	0.99	**0.37**	**0.79**	1.19	0.82	0.91	1.19	**0.59**
Threonine	0.83	0.96	0.56	1.05	0.80	**0.44**	0.46	0.58	**0.40**	**0.87**	**0.71**	0.33
Valine	**0.62**	1.08	0.74	1.12	1.15	**0.32**	**0.72**	**1.43**	0.91	1.37	**1.71**	0.70
Fructose	**1.20**	1.03	1.04	1.08	**0.82**	0.93	1.09	0.92	0.94	1.02	1.03	1.09
Fructose-6-P	1.22	1.14	0.61	0.88	nd	0.77	**1.27**	1.01	1.06	1.16	**0.69**	**1.93**
Gentiobiose	**1.97**	1.31	1.03	1.20	1.18	**0.70**	1.16	0.96	0.76	**1.21**	1.04	1.05
Glucose	**1.14**	0.96	0.96	**1.09**	**0.79**	1.02	1.28	0.90	**0.54**	1.49	**0.74**	0.82
Glucose-6-P	1.20	1.11	0.65	1.01	1.03	0.68	**1.45**	**1.31**	0.99	**1.30**	1.12	**1.42**
Isomaltose	**1.95**	1.32	1.31	1.27	1.18	1.07	1.81	0.85	0.62	1.41	**2.93**	1.16
Maltose	1.49	1.00	0.53	0.61	5.37	0.73	1.27	0.93	1.37	0.96	1.25	**1.50**
Sucrose	1.63	1.82	0.68	**2.44**	**1.82**	**2.25**	**1.73**	**1.24**	0.79	**1.68**	**1.38**	**1.77**
Trehalose	**1.46**	1.02	0.97	**1.39**	2.11	0.83	1.93	1.01	1.16	**2.01**	**6.39**	1.33
Xylose	0.97	0.94	1.43	0.95	0.69	0.79	**0.63**	**0.65**	**0.73**	**0.56**	**0.61**	0.79
Glycerol	1.25	1.15	1.09	1.00	0.82	1.38	1.24	1.47	0.65	0.86	0.58	0.67
Myo-ino-1-P	1.31	0.96	0.81	1.01	0.94	nd	1.29	1.16	0.94	0.93	1.09	0.97
Myo-ino	**0.70**	**0.69**	0.56	0.99	1.06	**0.38**	0.99	0.84	**0.72**	1.03	1.07	0.95
Benzoate	**1.87**	1.13	1.38	**2.26**	1.81	0.90	1.15	1.43	1.34	1.01	1.02	1.14
Citramalate	0.90	**0.73**	1.18	0.85	0.72	0.86	1.29	1.08	1.10	1.30	1.27	1.26
Citrate	**1.26**	1.04	0.89	**1.15**	**0.89**	**0.73**	1.08	0.93	0.88	0.98	0.95	1.04
Gluconate	1.17	1.07	1.17	**1.19**	1.33	**0.66**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.29**
Glycerate	**0.79**	**0.75**	0.78	**0.80**	0.74	**0.43**	0.90	**0.69**	1.40	0.98	1.09	1.03
Malate	**0.67**	0.96	0.67	**1.66**	1.23	**0.27**	1.09	0.92	0.97	**1.50**	1.12	0.73
Nicotinate	1.30	1.03	0.99	1.05	0.90	1.63	0.92	0.81	**0.74**	1.06	0.88	0.91
Phosphorate	**1.27**	0.96	0.97	1.01	0.77	**0.68**	0.96	0.90	1.10	0.92	1.12	0.90
Saccharate	**0.71**	**0.14**	0.47	**0.43**	0.21	**0.26**	0.95	0.78	0.79	0.90	0.89	1.07
Succinate	1.16	**1.16**	0.83	0.94	1.00	1.68	**1.71**	1.20	0.94	0.81	1.27	**1.91**
16:00	**1.20**	1.03	1.04	0.99	1.69	0.38	**0.59**	0.71	1.41	1.23	0.87	0.76
18:00	**1.19**	0.95	0.94	0.91	0.98	0.36	1.04	0.90	1.02	1.18	0.92	1.09

(E)-2-methyl-2-butenal, hexanal, (E)-2-heptenal, hexanoic acid, and acetophenone displayed transgressive behaviour in the B lines, whilst 1-penten-3-ol, pentanal, 2-ethylfuran, α-pinene, benzaldehyde, 1-nitro-2-phenylethane, β-damascenone, and geranylacetone exhibited such behaviour in the L lines. A total of 18 volatiles were transgressive in the B-derived lines, with a range of variation of 0.01–5.03 (ratio of relative abundance of the most extreme compounds compared with the parental line). Similarly, 15 volatiles were transgressive in the L lines, with a relative range of variation of between 0.01 and 12.8. Unlike the situation observed for primary metabolites, there is no a clear increase in the overall volatile content in the introgression lines. Indeed, the most remarkable differences are the dramatic decrease in a group of phenylpropanoid derivatives: eugenol, methylsalicylate, ethylsalicylate, and

Table 3. Metabolic analysis of the lines derived from the cross between Levovil and Cervil parents

Values are presented as the mean of six biologically independent determinations. The fold changes are relative to the Levovil parent. In bold are those values which were significantly different with $P < 0.05$ by the performance of Student's *t*-tests.

	Harvest 2004							Harvest 2005					
	L1	L2	L4	L9a	L9b	Lx	CxL	L1	L4	L9a	L9b	Lx	CxL
β-Alanine	2.26	2.49	1.31	2.24	1.44	0.75	**5.88**	1.53	**0.63**	1.56	0.85	0.84	0.86
Alanine	**9.58**	**7.52**	**3.22**	1.97	1.28	**4.42**	2.20	**2.58**	0.63	0.77	0.58	**2.52**	1.52
Asparagine	**3.31**	**2.78**	1.77	2.14	1.37	0.89	**3.00**	1.19	0.62	0.58	0.60	0.58	0.51
Aspartate	**2.25**	1.61	1.78	2.06	1.12	0.74	**2.68**	**1.97**	1.12	1.80	1.23	0.88	1.35
γ-Aminobutyrate	**2.02**	**2.29**	1.42	1.90	6.71	0.95	2.26	0.89	**0.48**	0.77	**0.60**	**0.51**	**0.40**
Glutamate	10.80	**14.83**	6.78	3.35	3.68	1.21	**3.80**	1.27	1.17	1.44	1.38	0.91	**2.07**
Glutamine	3.39	1.33	0.60	1.25	0.62	0.72	2.74	1.68	0.73	6.45	.63	0.67	0.62
Glycine	0.50	0.65	0.45	0.45	0.45	0.20	0.81	1.30	0.68	1.53	0.78	0.79	0.61
Homoserine	1.92	1.52	0.81	1.21	1.26	0.77	1.67	0.62	0.53	0.68	0.60	0.61	0.69
Isoleucine	**2.38**	**2.11**	1.17	1.58	1.64	0.88	**2.08**	**1.39**	0.83	1.18	1.30	0.88	1.12
Leucine	**3.59**	**2.76**	**1.91**	1.80	**1.97**	1.36	0.03	**1.67**	0.99	**1.34**	1.38	1.11	**1.56**
Lysine	4.49	4.77	3.01	0.92	1.38	0.73	2.02	1.35	0.69	1.58	0.83	1.41	**2.32**
Phenylalanine	**2.02**	1.92	1.36	1.39	**1.92**	0.84	2.08	1.24	0.61	0.62	0.79	0.64	0.64
Proline	**23.25**	**13.66**	7.15	**15.86**	1.30	13.05	nd	**2.75**	0.85	1.50	0.89	**1.66**	**6.01**
Putrescine	0.88	1.40	1.49	0.90	1.41	1.12	1.22	1.49	0.61	0.72	0.61	0.67	0.62
Pyroglutamate	**1.80**	**1.68**	1.35	**1.49**	1.31	0.99	**2.31**	1.08	0.68	0.89	0.65	0.66	0.64
Serine	1.71	1.56	1.14	1.73	1.29	0.63	1.01	**2.25**	0.86	**2.04**	1.57	0.85	0.81
Threonine	1.91	1.51	1.15	1.09	1.17	0.64	1.98	1.41	0.59	0.67	1.30	0.55	0.62
Tyrosine	2.54	2.92	1.53	0.56	1.45	0.22	1.40	**2.18**	**1.43**	**2.06**	**1.60**	**2.34**	**3.30**
Valine	1.54	1.60	1.64	1.44	1.13	0.46	**2.33**	1.37	0.88	0.74	0.75	0.66	0.73
Arabinose	**3.47**	**2.58**	1.16	1.60	1.24	1.31	1.23	1.68	0.73	6.45	0.63	0.67	0.62
Fructose	**1.61**	1.26	1.05	**1.48**	1.20	**1.71**	1.45	0.94	0.89	1.07	0.92	1.11	1.11
Fructose-6-P	**3.48**	**2.09**	1.56	**1.80**	1.45	**2.92**	**2.02**	**1.86**	1.30	**2.11**	1.36	**3.86**	**4.66**
Gentiobiose	**3.54**	2.00	1.28	1.67	2.61	**2.46**	**13.46**	1.16	0.90	**1.41**	1.04	1.27	**1.56**
Glucose	**1.35**	**1.27**	1.23	**1.32**	1.18	**1.34**	**1.49**	**0.79**	0.86	0.85	0.95	0.85	**0.66**
Glucose-6-P	**3.08**	1.80	1.39	**1.86**	1.27	**2.72**	**2.24**	**2.81**	**1.71**	**2.46**	**1.72**	**2.64**	**3.83**
Isomaltose	**4.72**	2.89	2.20	**4.43**	0.89	**4.00**	**2.43**	**3.32**	1.22	**1.95**	**1.80**	**2.49**	**5.54**
Maltose	**4.67**	2.11	1.67	**4.26**	nd	**2.91**	**2.33**	1.24	1.08	**2.04**	**1.24**	**2.46**	**3.97**
Rhamnose	**3.15**	**2.51**	0.93	1.17	1.11	0.94	1.06	**1.49**	0.88	1.15	1.07	1.38	**1.55**
Sucrose	**6.53**	**4.38**	2.34	**5.80**	**2.48**	**9.12**	**15.97**	**2.38**	**0.54**	**1.60**	0.73	**6.00**	**4.68**
Trehalose	**2.63**	**2.08**	1.72	**2.88**	1.17	**2.51**	**2.73**	**2.13**	**1.26**	**2.32**	**1.48**	**3.50**	**4.78**
Xylose	1.31	1.19	1.21	1.32	1.08	**1.46**	**0.25**	0.88	1.28	**1.49**	1.06	1.16	1.12
Glycerol	**1.48**	**1.52**	1.37	**1.32**	**1.44**	1.24	**13.73**	1.21	1.21	1.81	1.39	1.29	1.49
Myo-ino-1-P	**2.17**	**1.75**	**1.71**	1.21	1.33	**1.59**	1.15	**2.33**	**1.54**	1.28	1.15	0.86	0.87
Myo-ino	**1.56**	1.28	1.36	**1.82**	1.11	1.12	1.35	1.51	1.04	1.36	1.35	1.99	**2.62**
Benzoate	1.43	1.46	1.39	1.36	1.28	1.26	1.87	1.17	0.92	**1.37**	1.08	1.07	1.39
Citramalate	**2.34**	1.47	1.23	1.04	1.32	**2.05**	**2.04**	**2.52**	1.56	**1.55**	**2.17**	**3.26**	**3.61**
Citrate	**1.60**	**1.44**	1.18	**1.45**	1.16	**1.37**	**1.68**	**0.76**	0.93	0.76	1.07	**0.54**	**0.33**
Gluconate	**1.88**	1.50	1.14	1.42	0.89	**1.67**	**1.79**	1.38	1.20	**1.61**	**1.76**	**2.13**	**2.89**
Glycerate	0.96	1.06	1.02	0.89	1.48	1.18	**0.59**	**2.22**	1.98	1.19	1.34	1.22	**0.52**
Malate	0.76	0.79	0.83	1.05	1.02	**0.49**	**0.56**	**0.71**	0.89	**0.83**	0.83	**0.72**	**0.65**
Nicotinate	**2.40**	**2.01**	**1.75**	**1.74**	**1.43**	1.56	**35.60**	**1.59**	1.18	1.57	0.86	**1.59**	**1.98**
Phosphorate	**1.78**	**1.67**	1.14	**1.31**	1.23	1.11	1.50	**2.39**	1.30	**1.95**	1.23	1.20	1.09
Saccharate	**0.52**	0.79	**0.59**	1.16	**0.75**	**0.67**	**0.42**	1.44	0.86	1.37	1.42	1.66	**1.90**
Succinate	**2.58**	1.85	1.23	**1.56**	**1.38**	**3.94**	1.44	1.41	0.59	0.67	1.30	0.55	0.62
16:00	1.16	1.10	0.98	1.38	1.32	1.03	1.12	1.36	0.84	1.32	0.88	1.12	1.37
18:00	1.28	1.17	nd	1.36	1.32	1.06	1.08	**2.56**	**1.70**	**2.09**	**1.73**	**2.58**	**3.47**

guaiacol, to barely detectable levels in the lines harbouring a fragment of chromosome 9. The differences in the volatile patterns between the introgression lines and the varieties from which they are derived should thus be attributed more to the differences in levels of individual volatiles (or families thereof) rather than to differences in the overall volatile content.

Comparison of the levels of volatiles in the independent harvests (see Tables 4, 5) revealed that in contrast to the primary metabolite content, the data sets displayed large variation, indicating an important influence of environmental factors. The mean difference across the two harvests in the content of any given metabolite ranged

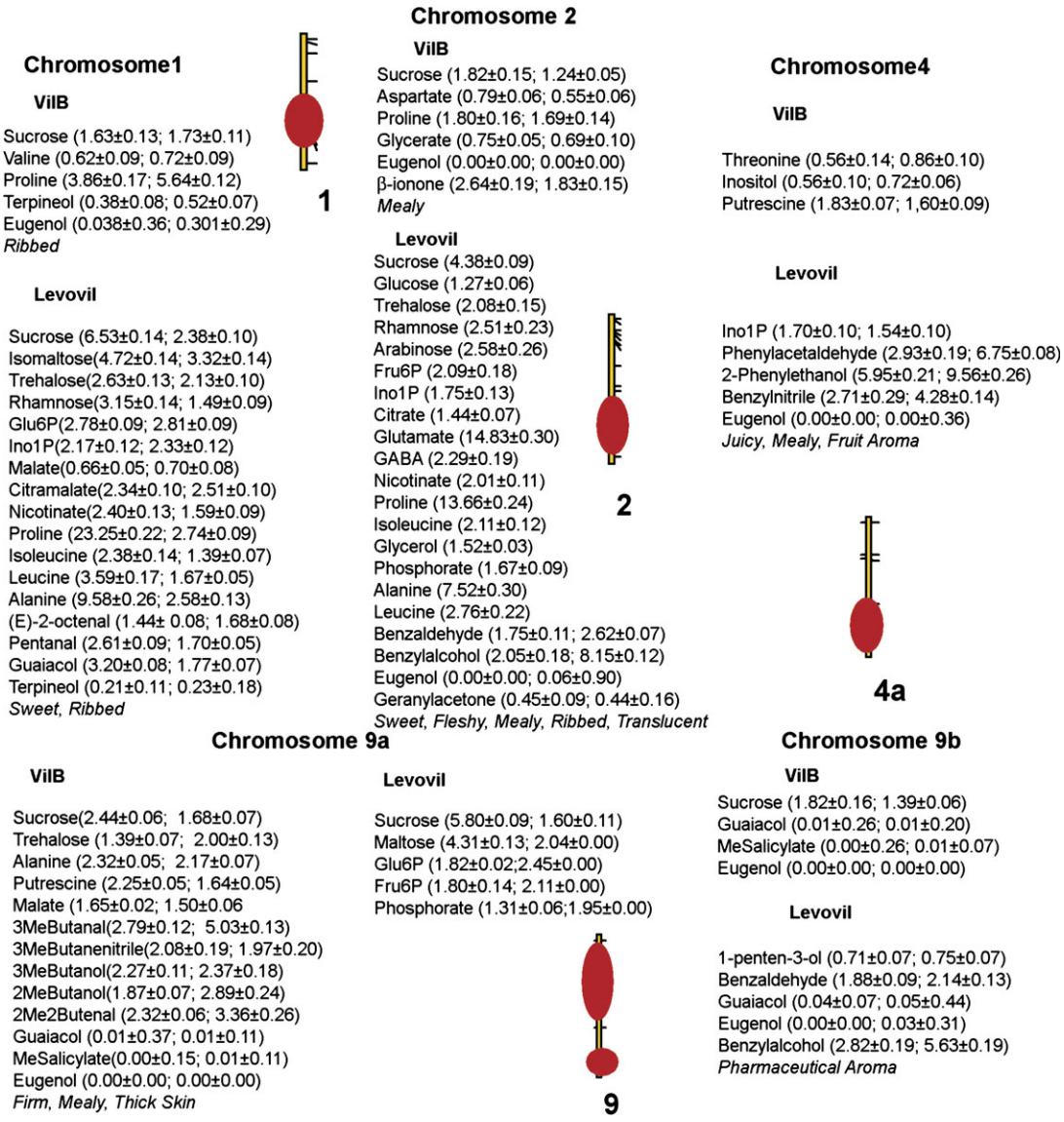

Fig. 2. Quantitative trait loci controlling the content of the primary metabolites, volatiles, and sensory properties (in italics) in VilB- and Levovil-derived lines. In parentheses are the fold changes relative to the respective parent (Lor B) for the two years (except for L2 in which only one year was analysed).

between 0.00 and 75.18 times the value observed in the L line for L recipient genotypes and between 0.00 and 79.12 times the value observed in the B line for B recipient genotypes.

QTLs were determined for these traits, revealing a total of 17 QTLs in the L background and 15 in the B background (see Fig. 2 and Tables 4, 5). Whilst many of the QTLs presented here were previously uncharacterized, several, including those for pentanal, (E)-2-methyl-2-butenal, guaiacol, and eugenol, have already been documented within this population (Saliba-Colombani *et al.*, 2001), whereas others, including 3-methylbutanal, 3-methylbutanenitrile, 3-methylbutanol, 2-methyl-1-butanol, and β-ionone, have also been previously described in the *S. pennellii* introgression lines (Tieman *et al.*, 2006*a*). The number of QTLs for volatiles was similar irrespective to the background into which the C genome segments were introgressed, with both L and B displaying approximately similar numbers of QTLs. Principal component analysis illustrates how many of the introgression lines are clearly distinguishable on the basis of their volatile profile. Variance in the levels of a group of phenolic derivatives (1-nitro-2-phenylethane, 2-phenylethanol, phenylacetaldehyde, and benzylnitrile) are responsible for the discrimination of the introgression line which harbours chromosome 4 fragments, whilst other NILs are segregated by their relative levels of other volatile compounds (Supplementary Fig. S1 at *JXB* online).

There are many co-localizations of volatile and organoleptic QTLs. The fruit aroma QTL co-localized with the QTL for 2-phenylethanol, benzylnitrile, and phenylacetaldehyde (chromosome 4), all of them phenolic derivatives with increased contents in the lines containing C alleles at

Table 4. List of volatiles measured on fruits harvested from VilB-derived lines

Values are presented as the mean of six biologically independent determinations. In bold are those values which were significantly different with $P < 0.05$ by the performance of Student's *t*-tests.

	Harvest 2005						Harvest 2006					
	B1	**B2**	**B4**	**B9a**	**B9b**	**Bx**	**B1**	**B2**	**B4**	**B9a**	**B9b**	**Bx**
2-Methyl-1-propanol	0.62	0.47	0.81	0.71	0.78	0.69	1.18	0.73	0.84	1.39	1.78	0.72
3-Methylbutanal	1.14	1.35	**0.20**	**2.80**	1.18	1.03	**2.55**	1.00	0.53	**5.03**	1.12	0.92
Butanol	1.07	0.89	0.81	1.60	1.20	1.27	0.94	1.48	0.58	1.98	3.75	1.56
1-Penten-3-ol	1.24	**1.25**	0.94	1.11	0.99	1.42	0.97	0.78	0.88	0.98	0.86	1.20
1-Penten-3-one	1.09	1.26	0.88	0.96	1.06	1.20	0.89	0.84	0.89	1.07	0.94	1.26
Pentanal	1.20	1.17	0.81	**0.64**	0.99	0.95	1.26	1.18	0.92	0.89	3.73	1.24
2-Ethylfuran	**1.78**	1.57	**1.68**	**2.12**	**1.72**	**2.24**	0.90	0.84	0.97	0.94	0.78	0.79
3-Methyl butanenitrile	0.94	1.65	**0.30**	**2.09**	2.55	1.51	**1.74**	1.73	0.46	**1.97**	1.57	0.93
3-Methylbutanol	1.38	1.18	**0.20**	**2.27**	1.32	1.08	**1.87**	0.80	0.71	**2.37**	**1.52**	0.97
2-Methyl-1-butanol	0.81	1.06	**0.40**	**1.87**	1.25	1.11	1.30	0.84	0.88	**2.89**	**2.29**	1.39
(E)-2-Methyl-2-butenal	1.25	0.88	0.65	**2.32**	1.32	0.92	1.14	0.51	0.66	**3.36**	2.65	0.96
(E)-2-Pentenal	1.05	**1.30**	0.87	0.91	1.11	5.02	0.80	0.74	0.74	0.92	0.80	1.20
1-Pentanol	**1.34**	1.01	0.91	0.94	1.09	0.99	1.16	0.90	0.90	0.86	0.89	1.16
(Z)-3-Hexenal	1.32	1.36	1.37	**1.63**	1.35	1.30	0.81	1.08	1.16	1.10	1.08	1.14
Hexanal	1.02	1.04	0.92	0.92	0.90	0.93	**1.38**	**1.30**	1.18	1.05	**1.27**	1.05
3-Methylbutanoic acid	2.70	1.79	0.51	**2.34**	1.73	1.35	**2.45**	0.49	0.49	1.71	1.52	0.70
(E)-2-Hexenal	1.30	1.38	1.20	**1.45**	1.04	**1.71**	0.85	0.83	0.96	0.92	0.84	0.94
(Z)-3-Hexen-1-ol	1.84	1.05	1.07	1.12	0.75	1.23	**1.48**	0.85	1.05	1.06	0.73	1.08
Pentanoic acid	0.82	0.80	0.62	**0.63**	**0.65**	**0.65**	1.26	0.44	0.94	0.92	1.11	0.93
(E,E)-2,4-Hexadien-1-al	1.03	0.98	1.37	1.32	1.21	1.49	0.84	1.01	1.04	0.99	1.02	1.07
α-Pinene	0.68	1.00	1.44	0.74	3.42	3.55	1.16	1.00	1.09	0.77	1.16	0.93
(E)-2-Heptenal	**1.36**	2.32	2.16	2.97	1.17	2.28	0.91	1.08	1.12	1.16	1.03	1.22
Benzaldehyde	0.62	1.42	0.98	0.51	1.05	**2.48**	0.73	1.02	0.71	1.01	0.89	**2.64**
6-Methyl-5-hepten-2-one	0.83	1.04	1.15	0.98	1.19	1.12	**1.73**	1.42	108	1.12	1.50	**1.67**
2-Pentylfuran	0.87	1.29	0.87	1.22	1.28	1.30	3.95	1.46	1.45	1.27	1.53	1.32
Hexanoic acid	0.60	0.70	0.58	0.62	1.02	0.74	1.08	1.41	1.29	4.96	0.70	1.10
Octanal	0.61	0.60	0.96	0.58	0.73	0.69	1.17	0.67	1.14	0.88	1.46	0.67
Benzylalcohol	**0.35**	1.23	0.92	**0.45**	1.07	**2.22**	0.95	1.04	0.79	0.63	1.50	**1.71**
2-Isobutylthiazole	0.78	0.81	0.27	0.67	0.83	0.61	**1.76**	1.24	0.41	0.86	0.87	0.77
Phenylacetaldehyde	0.48	1.30	0.64	0.73	0.72	1.10	0.70	1.46	**0.59**	1.23	0.76	1.43
(E)-2-Octenal	1.44	1.31	1.15	1.05	1.28	**1.66**	**1.81**	**1.37**	1.01	1.25	1.11	**2.36**
Acetophenone	0.60	0.85	1.23	0.65	0.57	0.81	0.79	1.26	1.04	0.86	0.81	1.01
p-Tolualdehyde	0.32	0.43	1.45	0.62	0.92	0.45	0.80	1.31	0.90	0.81	1.28	0.99
Guaiacol	**1.32**	**2.25**	**2.46**	**0.01**	**0.01**	**0.03**	0.74	0.68	1.54	**0.01**	**0.02**	**0.03**
Linalool	0.28	1.11	1.35	1.20	1.14	0.25	0.62	1.69	1.24	0.97	1.42	**0.43**
Nonanal	0.66	0.69	1.23	0.64	0.83	0.62	1.26	1.15	1.31	1.20	**1.57**	0.92
2-Phenylethanol	**0.65**	1.21	0.82	0.89	1.19	1.11	0.79	1.16	0.75	1.20	1.06	**1.60**
2-Ethyl-hexanoic acid	0.67	0.84	1.37	0.89	1.61	1.19	0.48	1.27	0.86	0.98	0.28	0.70
Benzylnitrile	0.52	1.31	0.97	0.79	1.47	1.13	1.19	**2.38**	0.83	1.98	1.77	2.38
Octanoic acid	0.65	0.77	1.42	0.85	1.34	1.01	0.08	0.52	0.28	0.34	0.04	0.27
Terpineol	**0.38**	0.97	**1.62**	0.95	1.03	**0.43**	**0.52**	**1.63**	1.21	0.90	**1.39**	**0.45**
Methyl salicylate	**2.18**	2.08	2.18	**0.00**	**0.00**	**0.01**	0.92	**0.39**	0.91	**0.01**	**0.01**	**0.01**
Geranial	0.75	0.85	1.55	0.92	1.21	1.16	2.14	2.05	2.08	2.23	1.74	1.67
Ethylsalicylate	0.49	**0.13**	1.62	**0.00**	**0.01**	0.01	2.97	2.06	3.43	0.00	0.74	1.34
1-Nitro-2-phenylethane	0.52	0.91	0.49	0.80	1.42	1.14	1.66	**3.19**	0.58	2.68	1.57	**5.11**
(E,E)-2,4-Decadienal	1.28	1.74	1.28	1.42	1.63	**2.66**	**2.32**	**1.40**	0.97	1.36	1.54	**2.90**
Eugenol	**0.04**	**0.00**	0.56	**0.00**	**0.00**	**0.00**	**0.30**	**0.00**	1.30	**0.00**	**0.00**	**0.00**
β-Damascenone	0.69	1.10	1.04	0.69	0.96	0.52	1.31	1.26	0.80	0.92	1.45	0.67
Geranylacetone	1.00	1.65	1.40	1.21	1.27	1.97	1.96	1.53	1.53	1.69	1.70	1.94
β-Ionone	1.99	**2.64**	1.82	**1.76**	**1.47**	**3.48**	**2.05**	**1.83**	1.50	2.16	1.69	**2.45**

this QTL. 2-Phenylethanol, the volatile which showed the highest increase, has been described to provide a sweet and fruity aroma (Togari *et al.*, 1995), and could be responsible for this fruit aroma perception. Pharmaceutical aroma QTL co-localized on chromosome 9 with the QTL of guaiacol and methylsalicylate, with both

Table 5. List of volatiles measured on fruits harvested from Levovil-derived lines

Values are presented as the mean of six biologically independent determinations. In bold are those values which were significantly different with $P < 0.05$ by the performance of Student's *t*-tests.

	Harvest 2005							Harvest 2006						
	L1	L2	L4	L9a	L9b	Lx	CxL	L1	L2	L4	L9a	L9b	Lx	CxL
2-Me-1-propanol	1.33	0.68	2.35	0.92	0.77	0.78	1.26	0.77	**0.46**	1.48	0.56	0.86	**0.23**	1.25
3-Methylbutanal	0.82	0.86	1.04	0.98	1.46	0.85	0.79	1.90	0.78	0.17	0.95	1.04	0.31	0.10
Butanol	1.80	**0.28**	1.31	0.70	1.30	0.53	1.26	2.15	0.71	1.34	1.01	2.13	0.58	4.04
1-Penten-3-ol	1.12	0.88	0.77	1.01	**0.71**	1.05	1.18	0.90	0.87	**0.77**	0.94	**0.75**	1.03	1.14
1-Penten-3-one	**1.60**	0.86	1.15	0.97	0.91	1.17	1.24	1.18	0.99	1.08	1.01	0.91	1.08	1.03
Pentanal	**2.61**	0.93	1.83	1.28	0.68	1.48	1.56	1.70	1.25	1.03	**0.77**	**0.70**	**1.36**	1.24
2-Ethylfuran	0.71	**1.61**	**0.57**	0.86	0.82	1.30	**1.45**	0.52	0.92	0.73	0.76	0.74	0.74	1.43
3-Me butanenitrile	**0.00**	**0.36**	0.60	0.75	**0.64**	**0.64**	**0.62**	1.29	1.22	0.57	1.32	0.69	0.78	0.50
3-Methylbutanol	0.81	0.94	0.91	1.00	1.48	0.86	1.34	0.62	0.70	0.52	0.99	1.24	0.57	1.11
2-Methyl-1-butanol	2.15	0.73	4.40	1.15	**2.81**	0.71	**1.78**	0.79	**0.48**	1.76	0.90	1.57	0.84	1.33
(E)-2-Me-2-butenal	0.99	0.52	12.8	0.80	1.16	0.56	**1.76**	0.28	0.90	**3.46**	0.75	0.78	0.83	1.50
(E)-2-Pentenal	**1.82**	0.87	1.27	1.11	0.91	1.36	**1.59**	1.08	1.28	1.19	1.03	0.78	1.15	1.40
1-Pentanol	1.36	0.84	1.19	1.24	0.74	1.19	1.39	1.47	1.16	0.97	0.77	0.72	1.30	**1.36**
(Z)-3-Hexenal	0.85	1.04	0.90	1.01	0.97	1.30	**1.47**	0.65	0.94	0.96	1.21	1.22	1.07	1.13
Hexanal	0.75	0.75	0.76	0.80	0.84	0.98	0.73	1.26	1.14	0.97	**0.82**	0.81	**1.25**	0.84
3-Me-butanoic acid	0.68	1.36	0.77	1.23	2.23	1.53	1.09	0.62	2.31	0.82	2.23	1.55	1.53	1.04
(E)-2-Hexenal	**1.57**	**1.56**	1.00	1.02	0.99	**1.29**	1.79	1.12	1.30	0.92	0.96	1.00	1.12	0.99
(Z)-3-Hexen-1-ol	**0.65**	0.75	0.73	1.11	**0.48**	**0.69**	0.79	0.71	1.16	1.16	1.33	0.98	1.29	**1.84**
Pentanoic acid	0.93	0.98	1.45	1.59	1.58	1.05	0.84	0.63	1.51	1.17	0.97	0.55	0.69	1.54
(E,E)-2,4-Hexadien-1-al	0.91	0.83	0.79	1.00	0.78	1.16	1.24	0.85	1.00	0.94	0.99	1.01	1.01	1.00
α-Pinene	**3.00**	**2.75**	23.6	9.39	1.04	1.85	1.80	0.65	0.89	1.01	0.86	13.18	**0.55**	0.76
(E)-2-Heptenal	**1.47**	1.27	1.40	**1.28**	0.85	2.47	1.15	0.67	0.89	0.84	1.15	1.04	1.07	1.32
Benzaldehyde	0.74	**1.75**	1.28	1.29	**1.88**	**1.99**	1.14	1.87	**2.62**	**3.78**	**3.19**	**2.15**	**4.07**	**2.29**
6-Me-5-hepten-2-one	0.35	**0.25**	0.55	0.43	0.58	0.44	0.35	0.90	0.66	0.63	**0.55**	0.56	0.69	**0.42**
2-Pentylfuran	0.79	0.79	0.96	0.97	0.99	1.32	0.75	0.96	0.92	0.76	0.79	0.60	1.09	0.85
Hexanoic acid	2.29	1.32	1.71	1.67	1.99	1.41	0.91	0.88	2.13	2.22	2.77	1.97	1.57	2.66
Octanal	1.03	0.96	1.01	1.11	0.93	0.77	0.59	1.08	0.98	0.74	0.67	0.78	0.61	0.40
Benzylalcohol	0.95	**2.05**	1.09	1.32	**2.82**	**2.68**	1.02	8.15	**7.00**	**9.84**	9.15	**5.63**	**16.6**	**4.78**
2-Isobutylthiazole	**0.15**	**0.29**	**0.36**	**0.08**	**0.28**	**0.08**	**0.10**	0.88	1.22	0.51	0.12	0.50	**0.19**	**0.20**
Phenylacetaldehyde	0.99	**0.54**	**2.93**	0.62	0.75	0.74	1.43	1.77	1.73	**6.76**	1.02	1.52	1.38	**4.77**
(E)-2-Octenal	**1.44**	1.00	0.97	1.25	0.71	**1.63**	**1.47**	1.68	1.21	0.85	**0.91**	**0.71**	**2.10**	**2.36**
Acetophenone	0.80	0.58	0.62	0.75	0.56	0.71	0.56	0.83	0.83	0.93	0.76	1.10	0.72	1.06
p-Tolualdehyde	0.79	1.36	2.14	1.73	1.56	**0.56**	0.93	0.95	0.55	0.45	0.60	1.00	0.68	2.37
Guaiacol	**3.20**	1.18	0.82	1.29	**0.04**	**0.02**	**0.01**	1.77	0.96	0.75	**0.83**	**0.05**	**0.02**	**0.01**
Linalool	**0.10**	0.55	0.83	0.57	0.77	**0.16**	**0.32**	0.23	1.05	0.83	1.35	1.09	0.24	0.45
Nonanal	1.11	0.96	1.14	1.06	1.11	0.65	0.68	085	0.84	0.75	0.97	0.83	**0.49**	0.60
2-Phenylethanol	0.67	0.73	**5.95**	**0.51**	0.88	0.59	**2.35**	1.90	1.18	**9.56**	0.93	1.46	**2.13**	7.56
2-Ethyl-hexanoic acid	1.74	1.36	1.97	1.28	1.82	1.46	0.77	1.53	3.08	3.57	5.21	2.14	1.89	3.28
Benzylnitrile	0.46	**0.34**	**2.71**	0.50	**0.42**	**0.43**	0.81	1.59	1.48	**4.38**	0.87	1.16	0.90	**2.31**
Octanoic acid	0.95	1.15	0.93	1.12	1.55	1.35	0.79	3.82	6.65	5.62	32.3	8.57	9.39	**16.66**
Terpineol	**0.21**	**0.55**	1.08	0.61	0.60	**0.16**	**0.34**	0.23	0.80	0.78	0.92	0.82	**0.18**	**0.40**
Methyl salicylate	**2.59**	**5.76**	0.68	3.37	**0.02**	**0.02**	**0.01**	1.36	1.64	1.17	1.26	0.09	0.07	0.08
Geranial	**0.37**	**0.35**	0.74	0.50	0.61	**0.47**	**0.33**	0.74	0.57	0.51	**0.43**	0.61	0.57	**0.39**
Ethylsalicylate	3.21	**0.19**	0.53	1.35	**0.00**	**0.00**	**0.00**	2.43	2.64	0.53	1.09	0.00	0.00	0.00
1-Nitro-2-phenylethane	0.43	**0.20**	5.84	0.69	0.47	**0.46**	0.65	1.89	1.08	**7.53**	0.59	0.70	1.27	**2.66**
(E,E)-2,4-Decadienal	1.07	0.88	1.06	1.60	0.53	1.65	**2.33**	1.29	0.80	0.69	0.54	**0.52**	1.72	1.51
Eugenol	1.13	**0.00**	**0.00**	0.70	**0.00**	**0.00**	**0.00**	0.62	**0.06**	**0.00**	0.60	**0.03**	**0.00**	**0.00**
β-Damascenone	1.13	1.29	1.42	1.12	0.73	0.44	1.11	0.70	0.77	0.50	1.20	1.50	0.56	0.53
Geranylacetone	0.72	**0.45**	0.77	0.64	0.66	0.81	0.76	0.72	**0.44**	**0.37**	**0.35**	0.53	0.87	**0.38**
β-Ionone	0.65	0.66	0.63	0.94	0.59	**1.78**	0.85	0.71	0.56	**0.45**	**0.51**	0.53	1.24	0.53

phenylpropanoid derivatives levels being ~20-fold lower in the lines containing the C alleles at this QTL. As previously stated, guaiacol and eugenol provide a medicinal-like aroma. Thus, these compounds could conceivably be responsible for the pharmaceutical aroma perception.

Correlation analysis

For a fuller characterization of the associations between traits, a correlation-based approach was adopted in which the mean values determined above for each metabolite were compared with those determined for each volatile. For this purpose, a combinatorial analysis of all metabolites (both primary and volatile) was carried out, by running the data points through pairwise correlation analysis. Of the 4560 possible pairs analysed, 806 and 750 resulted in significant correlations ($P \leqslant 0.05$) for L and B lines, respectively. Of these pairs, 609 and 466 showed positive ($r > 0.65$) and 197 and 284 showed negative (r less than -0.65) correlation coefficients for L- and B-derived lines, respectively. The heat map of Fig. 3 (and Supplementary Tables S2, S3 at JXB online) shows the correlations between primary metabolites and volatiles (to simplify interpretation, metabolites are grouped on the basis of their compound class). Negative correlations were significant between the sugars and sugar derivatives fructose, fructose-6P, glucose, glucose-6P, isomaltose, and sucrose, and the volatiles linalool, terpineol, and nonanal in both genetic backgrounds, whilst geranial was also strongly negatively correlated with sugars in the L background but not in the B background. In contrast, positive correlations were observed between 1-penten-3-ol, (E)-2-hexenal, (E)-2-octenal, and (E,E)-2,4-decadienal and the above-mentioned sugars. There is little correlation between the levels of the volatile organic compounds and their direct precursors from primary metabolism. Correlations within primary metabolites and volatiles were also analysed. The full data set of correlation coefficients is presented in Supplementary Tables S2 –S7. Among primary metabolites (Supplementary Tables S4, S5), correlations were qualitatively similar to those reported previously in data sets wherein metabolite contents varied either across a developmental time course (Carrari *et al.*, 2006) or across the

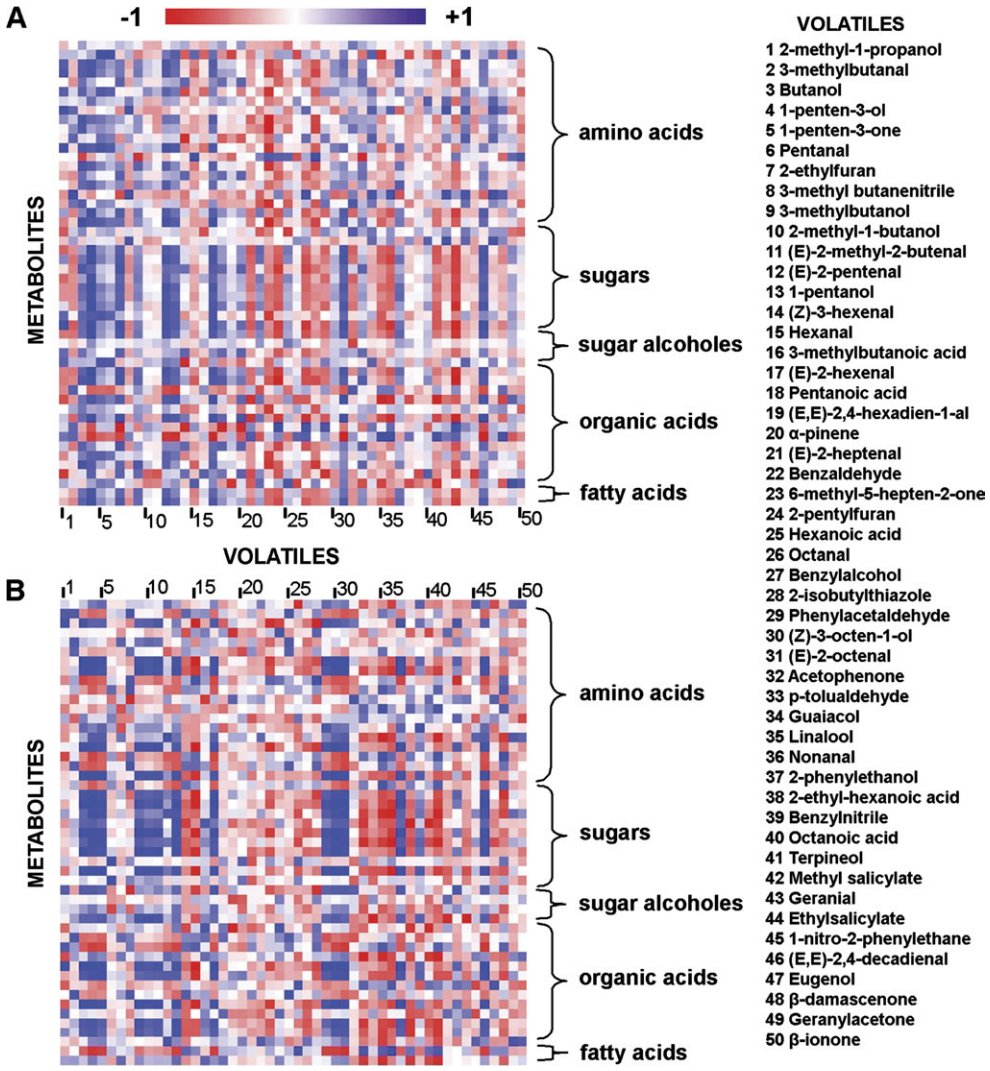

Fig. 3. Heat map showing the correlation analysis between traits in tomato NILs. (A) Mean of metabolites and volatiles during the two years for Levovil-derived NILs. (B) Mean of metabolites and volatiles during the two years for VilB-derived NILs. Regions in red and blue indicate negative or positive correlations between traits, respectively (the complete data set is also available in Supplementary Tables S2 and S3 at *JXB* online).

S. pennellii introgression lines (Schauer *et al.*, 2006, 2008). As observed previously in Carrari *et al.* (2006), phosphorylated intermediates displayed the greatest number of significant correlations to other primary metabolites. Among the different classes of primary metabolites, the sugars displayed the highest number of correlations irrespective to the genotype analysed. For example, sucrose, fructose, and glucose exhibited 20, 20, and 15 significant correlations in L-derived lines and 17, 21, and 19 in B-derived lines, respectively. Other compounds displayed a different number of correlations when the two genotypes were considered. Aspartate and asparagine displayed 23 and 21 significant correlations, respectively, in the L-derived lines but no significant correlations in the B-derived lines. Additionally the number of correlations for glutamate in the L-derived lines was lower compared with those observed in B-derived lines (10 and 15, respectively). γ-Aminobutyric acid (GABA) and saccharate displayed a low number of correlations in L-derived lines (0 and 3, respectively) but a high number in B-derived lines (12 and 15, respectively). Similarly, the volatile–volatile correlations (Supplementary Tables S6, S7) observed across the lines were largely in accordance with those described by Tikunov *et al.* (2005) across a panel of 94 tomato cultivars. The results were consistent with most of the previously described correlations such as those of eugenol, guaiacol, methylsalicylate, and ethylsalicylate. Some novel correlations were also uncovered in the present study such as those between 1-nitro-2-phenylethane and benzylnitrile or other phenylpropanoid derivatives, or the tight correlations between (E)-2-octenal and (E,E)-2,4-decadienal, or 1-penten-3-ol and other lipid derivatives. A strong correlation was additionally observed between linalool and terpineol, and also between 2-methyl-1-propanol, 2-phenylethanol, and butanol. As described for the primary metabolites, many of the correlations were observed in both genetic backgrounds (L and B), whilst others were significant only in one of them (Supplementary Tables S6, S7).

As a final analysis, correlations between all chemical traits measured in L-derived lines with organoleptic properties assessed on the same harvest were studied by sensory profiling (Fig. 4 and Supplementary Tables S8, S9 at *JXB* online). Of 1615 pairs of traits, 181 showed significant correlations ($P \leqslant 0.05$), among which 101 exhibited positive correlations ($r > 0.65$) and 80 displayed negative correlations (r less than –0.65). Some of the chemical traits showed opposite behaviour with respect to different sensory properties. For example, xylose correlated positively with firmness but negatively with juiciness, whilst malate correlated positively with sourness and negatively with sweetness. However, there were other cases, such as those of sweetness and global aroma, in which sensory traits displayed highly similar correlative behaviour with the same metabolites. When analysed specifically from the perspective of the organoleptic traits, some strong correlations were observed, such as colour intensity–glutamic acid (r=0.98), pharmaceutical aroma–guaiacol (r=0.97), typical tomato aroma–phenylalanine (r=–0.97), global aroma–2-ethyl-hexanoic acid [r=–0.98; global aroma corresponded to the general impression of aroma before swallowing (Causse *et al.*, 2001)], sweetness–citramalic acid (r=0.99), sourness–alanine (r=–0.97), juiciness–trehalose (r=–0.99), firmness–glutamic acid (r=0.99), embarrassing skin–xylose [r=–0.97; embarrasing skin is a sensory attribute which describes how difficult it is to swallow fruit skin and therefore it has a higher tendency to remain in the mouth (Causse *et al.*, 2001)]. Some of these correlations could probably be predicted on the basis of the chemical properties of the metabolites, such as, for example, the volatile guaiacol (which correlated positively with pharmaceutical aroma), is described as having a smoke-like or medicinal odour, and 2-ethyl-hexanoic acid (which negatively correlated with global aroma) which exhibited a wine-like odour. A more in-depth analysis of the organoleptic traits revealed complex interactions among many metabolites. Global aroma, for instance, significantly correlated to many volatiles, both positively [1-pentanol (r=1.00), (E)-2-hexenal (r=0.97), (E)-2-pentenal (r=0.93), 1-penten-3-one (r=0.91)] and negatively [2-ethylhexanoic acid (r=–0.98), pentanoic acid (r=–0.96), linalool (r=–0.95)], and also to non-volatile compounds [alanine (r=0.98)]. Typical tomato aroma displayed significant positive correlation only with the volatile benzaldehyde (r=0.91), but exhibits negative correlation with 12 metabolites, most of them being non-volatile.

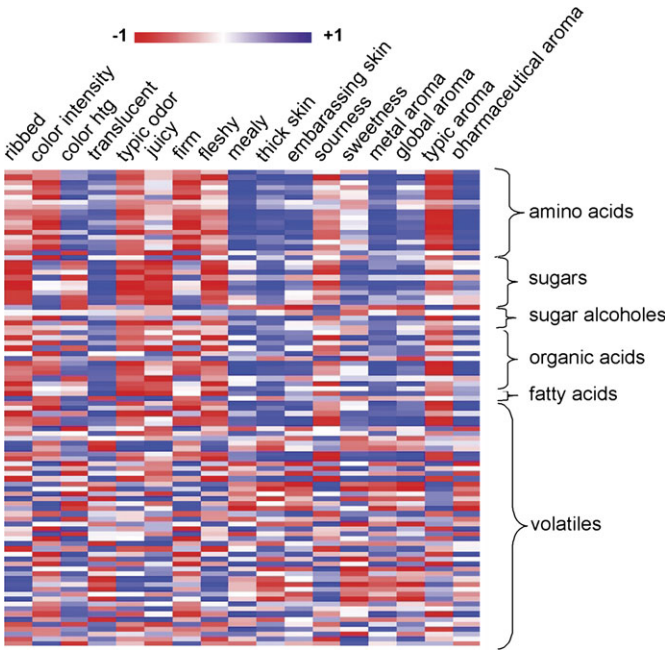

Fig. 4. Heat map showing the correlation analysis between primary metabolites, volatiles, and sensory properties in Levovil-derived tomato NILs. Regions in red and blue indicate negative or positive correlation between traits, respecively (for details see Supplementary Table S9 at *JXB* online).

Discussion

Fruit flavour is known to be considerably influenced by several factors. For example, the contents of primary

metabolites such as organic acids and sugars are known to be important, but the sugar/acid ratio is also an important determinant of taste. In practical terms, this can be summarized as follows: both high sugar content and acidity result in a good flavour, low acidity and high sugar content gives a bland flavour, high acidity and low sugar content give a tart flavour, and finally low acidity and sugar content results in an essentially tasteless flavour. On the other hand, volatile components which build fruit aroma greatly influence human perception of flavour. Here the metabolomic approach was used to describe the phenotypic variation of a broad range of primary and volatile metabolites across diverse genetic backgrounds. The results of the most highly abundant primary metabolite analysis of cherry and large-fruited tomatoes lines were largely in accordance with those obtained from previous studies (Causse *et al.*, 2002). The low sugar and high malate content of the L parent and the corresponding very low sugar/acid ratio could explain the lower acceptance of the fruit by the food panel tasters, especially given that malate is perceived as sourer tasting than citrate (Marsh *et al.*, 2003).

Other less abundant primary metabolites were also found at different levels in the parental lines. A recent survey of metabolite content in the fruits of a range of wild tomato species revealed that whilst these displayed large variations in sugar and amino acid content they were essentially unaltered in the content of tricarboxylic acid (TCA) cycle intermediates (Schauer *et al.*, 2005*b*). This suggests that the variation observed here is probably the result of breeding-based selection. One metabolite of particular interest is glutamate, known to be sensed as the fifth basic taste (umami), which evokes a savoury feeling. In addition to the changes observed in sugars and acids in cherry tomatoes, the glutamate level was found to be considerably higher in the C variety than in the large-fruited varieties. This finding is additionally in accordance with the fact that cherry tomatoes were found to be tastier than the other parental lines used in this study.

Within the aroma components, 2-phenylethanol is known to provide a sweet and fruity perception (Togari *et al.*, 1995). It is thus expected that the increased levels of 2-phenylethanol in line C would synergistically interact with sugars to produce an even sweeter flavour. Moreover, guaiacol has been described as an undesirable compound in many fruits, as it provides a medicinal-like aroma (Zierler *et al.*, 2004).

The evaluation of the primary metabolite content of a subset of tomato lines containing marker-defined introgressions, of five regions controlling fruit quality variation from the cherry tomato into large-fruited genetic backgrounds, revealed only a relatively small number of metabolites which exhibited transgressive behaviours across both harvests. This contrasted with the situation observed in interspecific introgression lines in which segments of the *S. pennellii* genome were inserted into the background of the M82 cultivar of *S. lycopersicum*, in which transgressive behaviour was observed for the majority of metabolic traits (Schauer *et al.*, 2006). Irrespective of whether they were transgressive or not, the changes in metabolites showed a strong bias toward an increase in metabolite contents in the introgressed lines relative to either recipient background. This could have been anticipated since the cherry tomato line was characterized as generally displaying a higher metabolite content than the large-fruited cultivars, but this is not true for all metabolites since increases were also found in the metabolite valine that was present at lower levels within the cherry tomato than in the large-fruited species.

As stated before, unlike the situation observed in primary metabolites, there is no clear increase in the overall volatile content of the introgression lines. Thus, the differences in the volatile pattern between parental and introgression lines are due to the differences in individual volatiles (or families of them), their modified levels depending on the introgressed chromosome fragment.

Few clear patterns emerged when co-localization between metabolite and volatile traits was examined. Co-localizations of QTLs for two metabolites could be due either to physiological relationships or to the action of two genes genetically linked and introgressed in the same region, as the size of introgressed regions is still large (~10–40 cM). For example, the negative association between sucrose and eugenol content must be due to genetic linkage rather than to a common physiological origin since there are other examples of these traits varying independently of one another and, moreover, the molecular mechanism underlying this association cannot be formally resolved in the current study. Evaluation of the *S. pennellii* introgression lines revealed that increased levels of 2-phenylethanol and 2-phenylacetaldehyde were independent of changes in the level of phenylalanine (Tieman *et al.*, 2006*b*). In this study, the content in volatiles correlated more with the levels of soluble sugars than with their direct precursors. The most likely explanation is that sink strength regulates part of the production of secondary metabolites. Nevertheless, it is also possible to speculate that these changes could be due to sugar-mediated changes in gene expression of enzymes involved in their biosynthetic pathways or that they merely resulted from spurious associations resulting from gene linkages within the large introgressions of the C genome. Considerably more experimental evidence is, however, required in order to provide mechanistic insight into these phenomena. This is indeed the case for any of the associations presented here since the data provided only indicate linkages between the various traits and do not provide any information concerning the causality underlying their association. Whilst some of the correlations found in the present work could probably be predicted on the basis of the chemical properties of the metabolites, the vast majority are novel, and as such could provide valuable information in helping to unravel the complex basis of sensory fruit traits. It seems likely that considerable research effort is still needed in order to identify the causality, if any, underlying these relationships.

Conclusion

A comprehensive profiling of both small molecule primary metabolites and the important volatile organic compounds of tomato was performed in independent cultivars of

tomato containing equivalent introgression regions from a cherry tomato variety. The results confirmed and extended earlier studies (Causse *et al.*, 2001, 2002, 2004), suggesting that chemical composition QTLs were identifiable and hence probably tractable from these crosses. In addition, they revealed that the expression of the QTLs is highly dependent on the genetic background, D-derived lines displaying far fewer QTLs for primary metabolites than L- and B-derived lines (a fact exacerbated when it is taken into account that the QTLs for the D genotype could only be regarded as putative). The current study utilized a broad level profiling of primary metabolites and volatiles to facilitate the evaluation of possible links between them. The lack of correlation between the levels of specific volatile organic compounds and the levels of their precursor metabolites is perhaps at first sight surprising. However, this is not without precedent since the levels of 2-phenylacetaldehyde and 2-phenylethanol have previously been shown to vary greatly independently of the levels of phenylalanine (Tieman *et al.*, 2006). This finding suggests that the rate of volatile production is generally not governed by precursor supply but rather at the transcriptional or post-transcriptional level. Although more studies will be required to understand the complex factors underlying consumer preference in tomato, the results provide several candidate molecules that may be useful leads for this purpose.

Supplementary data

Supplementary data are available in *JXB* online.

Figure S1 PCA analysis of volatiles.

Table S1 Primary metabolites in VilD-derived NILs.

Table S2–S3 Correlation analysis between primary metabolites and volatiles.

Table S4–S5 Correlation analysis between primary metabolites.

Table S6–S7 Correlation analysis between volatiles.

Table S8 Sensory profiling of Levovil-derived lines.

Table S9 Correlation analysis between primary metabolites, volatiles, and sensory properties in Levovil-derived lines.

Acknowledgements

The help of Nicolas Schauer (metabolite profiling), Aaron Fait (bioinformatic analysis), Emmanuel Botton and Yolanda Carretero (fruit harvests and sample preparations), Karine Robini (sensory data management), Cristina Alfaro and Jaime Primo (volatile analysis) is gratefully acknowledged. Furthermore, we acknowledge the financial support of the trilateral project GENMETFRUQUAL (BMBF FKZ 0313151).

References

Baldwin EA, Scott JW, Shewmaker CK, Schuch W. 2000. Flavor trivia and tomato aroma: biochemistry and possible mechanisms for control of important aroma components. *Hortscience* **35,** 1013–1022.

Brummell DA, Harpster MH. 2001. Cell wall metabolism in fruit softening and quality and its manipulation in transgenic plants. *Plant Molecular Biology* **47,** 311–340.

Buttery RG, Seifert RM, Guadagni DG, Ling LC. 1971. Characterization of additional volatile components of tomato. *Journal of Agricultural and Food Chemistry* **19,** 524–529.

Carrari F, Baxter B, Usadel B, *et al*. 2006. Integrated analysis of metabolite and transcript levels reveals the metabolic shifts that underlie tomato fruit development and highlight regulatory aspects of metabolic network behavior. *Plant Physiology* **142,** 1380–1396.

Causse M, Buret M, Robini K, Verschave P. 2003. Inheritance of nutritional and sensory quality traits in fresh market tomato and relation to consumer preferences. *Journal of Food Science* **68**(7), 2342–2350.

Causse M, Duffe P, Gomez MC, Buret M, Damidaux R, Zamir D, Gur A, Chevalier C, Lemaire-Chamley M, Rothan C. 2004. A genetic map of candidate genes and QTLs involved in tomato fruit size and composition. *Journal of Experimental Botany* **55,** 1671–1685.

Causse M, Saliba-Colombani V, Lecomte L, Duffé P, Rousselle P, Buret M. 2002. QTL analysis of fruit quality in fresh market tomato: a few chromosome regions control the variation of sensory and instrumental traits. *Journal of Experimental Botany* **53,** 2089–2098.

Causse M, Saliba-Colombani V, Lesschaeve I, Buret M. 2001. Genetic analysis of organoleptic quality in fresh market tomato. 2. Mapping QTLs for sensory attributes. *Theoretical and Applied Genetics* **102,** 273–283.

Chaïb J, Devaux MF, Grotte MG, Robini K, Causse M, Lahaye M, Marty I. 2007. Physiological relationships among physical, sensory, and morphological attributes of texture in tomato fruits. *Journal of Experimental Botany* **58,** 1915–1925.

Chaïb J, Lecomte L, Buret M, Causse M. 2006. Stability over genetic backgrounds, generations and years of quantitative trait locus (QTLs) for organoleptic quality in tomato. *Theoretical and Applied Genetics* **112,** 934–944.

Decoene C. 1995. Tomates, qu'en pensent les consommateurs? *Infos-Ctifl* **112,** 8–11.

DeGiglio MA. 2003. Growth of the fresh greenhouse tomato market in the USA. *Acta Horticulturae* **611,** 91–92.

Dennison RA, Hall CB, Nettles VF. 1953. Factors influencing tomato quality. *Proceedings of the Florida State Horticultural Society* **65,** 108–111.

Fernie AR, Tadmor Y, Zamir D. 2006. Natural genetic variation for improving crop quality. *Current Opinion in Plant Biology* **9,** 196–202.

Fernie AR, Trethewey RN, Krotzky AJ, Willmitzer L. 2004. Metabolite profiling: from diagnostics to systems biology. *Nature Reviews Molecular and Cellular Biology* **5,** 763–769.

Goff SA, Klee HJ. 2006. Plant volatile compounds: sensory cues for health and nutritional value? *Science* **311,** 815–819.

Harker FR, Stec MGH, Hallett IC, Bennet CL. 1997. Texture of parenchymatous plant tissue: a comparison between tensile and other instrumental and sensory measurements of tissue strength and juiciness. *Postharvest Biology and Technology* **11,** 63–72.

Hovav R, Chehanovsky N, Moy M, Jetter R, Schaffer AA. 2007. The identification of a gene (Cwp1), silenced during Solanum evolution, which causes cuticle microfissuring and dehydration when expressed in tomato fruit. *The Plant Journal* **52,** 627–639.

Janse J, Schols M. 1995. Une préférence pour un goût sucré et non farineux. *Groenten+Fruit* **26,** 16–17.

Kopka J, Schauer N, Krueger S, *et al*. 2005. GMD@CSB.DB: the Golm Metabolome Database. *Bioinformatics* **21,** 1635–1638.

Lecomte L, Duffé P, Buret M, Servin B, Hospital F, Causse M. 2004. Marker-assisted introgression of five QTLs controlling fruit quality traits into three tomato lines revealed interactions between QTLs and genetic backgrounds. *Theoretical and Applied Genetics* **109,** 658–668.

Lisec J, Meyer RC, Steinfath M, *et al*. 2008. Identification of metabolic and biomass QTL in Arabidopsis thaliana in a parallel analysis of RIL and IL populations. *The Plant Journal* **53,** 960–972.

Lisec J, Schauer N, Kopka J, Willmitzer L, Fernie AR. 2006. Gas chromatography mass spectrometry-based metabolite profiling in plants. *Nature Protocols* **1,** 387–396.

Marsh KB, Rossiter K, Lau K, Walker S, Gunson A, MacRae E. 2003. Using fruit pulps to explore flavour in kiwifruit. *Acta Horticulturae* **610,** 229–238.

Petro-Turza M. 1987. Flavour of tomato and tomato products. *Food Review International* **2,** 309–351.

Ratanachinakorn B, Klieber A, Simons DH. 1997. Effect of short-term controlled atmospheres and maturity on ripening and eating quality of tomatoes. *Postharvest Biology and Technology* **11,** 149–154.

Redgwell RJ, Fischer M. 2002. Fruit texture, cell wall metabolism and consumer perceptions. In: Knee M, ed. *Fruit quality and its biological basis*. Sheffield: Sheffield Academic Press, 46–88.

Roessner-Tunali U, Hegeman B, Lytovchenko A, Carrari F, Bruedigam C, Granot D, Fernie AR. 2003. Metabolic profiling of transgenic tomato plants overexpressing hexokinase reveals that the influence of hexose phosphorylation diminishes during fruit development. *Plant Physiology* **133,** 84–89.

Rowe HC, Hansen BG, Halkier BA, Kliebenstein DJ. 2008. Biochemical networks and epistasis shape the Arabidopsis thaliana metabolome. *The Plant Cell* **20,** 1191–216.

Saladie M, Matas AJ, Isaacson T, *et al*. 2007. A reevaluation of the key factors that influence tomato fruit softening and integrity. *Plant Physiology* **144,** 1012–1028.

Saliba-Colombani V, Causse M, Langlois D, Philouze J, Buret M. 2001. Genetic analysis of organoleptic quality in fresh market tomato. 1. Mapping QTLs for physical and chemical traits. *Theoretical and Applied Genetics* **102,** 259–272.

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

Schauer N, Semel Y, Balbo I, Steinfath M, Repsilber D, Selbig J, Pleban T, Zamir D, Fernie AR. 2008. Mode of inheritance of primary metabolic traits in tomato. *The Plant Cell* **20,** 509–523.

Schauer N, Semel Y, Roessner U, *et al*. 2006. Comprehensive metabolic profiling and phenotyping of interspecific introgression lines for tomato improvement. *Nature Biotechnology* **24,** 447–454.

Schauer N, Steinhauser D, Strelkov S, *et al*. 2005*a*. GC-MS libraries for the rapid identification of metabolites in complex biological samples. *FEBS Letters* **579,** 1332–1337.

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

Serrano-Megias M, Lopez-Nicolas JM. 2006. Application of agglomerative hierarchical clustering to identify consumer tomato preferences: influence of physicochemical and sensory characteristics on consumer response. *Journal of the Science of Food and Agriculture* **86,** 493–499.

Spencer JPE, Kuhnle GGC, Hajirezaei MR, Mock HP, Sonnewald U, Rice-Evans C. 2005. The genotypic variation of the antioxidant potential of different tomato varieties. *Free Radical Research* **39,** 1005–1016.

Stevens MA. 1972. Relationships between components contributing to quality variation among tomato lines. *Journal of the American Society for Horticultural Science* **97,** 70–73.

Szczesniak AS. 2002. Texture is a sensory property. *Food Quality and Preference* **13,** 215–225.

Tieman D, Taylor M, Schauer N, Fernie AR, Hanson AD, Klee HJ. 2006*b*. Tomato aromatic amino acid decarboxylases participate in synthesis of the flavor volatiles 2-phenylethanol and 2-phenylacetaldehyde. *Proceedings of the National Academy of Sciences, USA* **103,** 8287–8292.

Tieman DM, Zeigler M, Schmelz EA, Taylor MG, Bliss P, Kirst M, Klee HJ. 2006*a*. Identification of loci affecting flavour volatile emissions in tomato fruits. *Journal of Experimental Botany* **57,** 887–896.

Tikunov Y, Lommen A, de Vos CHR, Verhoeven HA, Bino RJ, Hall RD, Bovy AG. 2005. A novel approach for nontargeted data analysis for metabolomics. Large-scale profiling of tomato fruit volatiles *Plant Physiology* **139,** 1125–1137.

Togari N, Kobayashi A, Aishima T. 1995. Relating sensory properties of tea aroma to gas chromatographic data by chemometric calibration methods. *Food Research International* **28,** 485–493.

Zierler B, Siegmund B, Pfannhauser W. 2004. Determination of off-flavour compounds in apple juice caused by microorganisms using headspace solid phase microextraction-gas chromatography-mass spectrometry. *Analytica Chimica Acta* **520,** 3–11.