Journal of Experimental Botany, Vol. 60, No. 12, pp. 3379–3386, 2009
doi:10.1093/jxb/erp177 Advance Access publication 10 June, 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

Use of network analysis to capture key traits affecting tomato organoleptic quality

Paola Carli[1], Serena Arima[2], Vincenzo Fogliano[3], Luca Tardella[2], Luigi Frusciante[1] and Maria R. Ercolano[1,*]

[1] Department of Soil, Plant and Environmental and Animal Production Sciences, University of Naples 'Federico II', Via Università 100, 80055 Portici (NA), Italy
[2] Department of Statistics, Probability and Applied Statistics, University of Rome 'La Sapienza', P.le Aldo Moro 5, 00185 Rome, Italy
[3] Department of Food Science, University of Naples 'Federico II', Via Università 100, 80055 Portici (NA), Italy

Received 20 March 2009; Revised 30 April 2009; Accepted 6 May 2009

Abstract

The long-term objective of tomato breeders is to identify metabolites that contribute to defining the target flavour and to design strategies to enhance it. This paper reports the results of network analysis, based on metabolic phenotypic and sensory data, to highlight important relationships among such traits. This tool allowed a reduction in data set complexity, building a network consisting of 35 nodes and 74 links corresponding to the 74 significant (positive or negative) correlations among the variables studied. A number of links among traits contributing to fruit organoleptic quality and to the perception of sensory attributes were identified. Modular partitioning of the characteristics involved in fruit organoleptic perception captured the essential fruit parameters that regulate interactions among different class traits. The main feature of the network was the presence of three nodes interconnected among themselves (dry matter, pH, and °Brix) and with other traits, and nodes with widely different linkage degrees. Identification of strong associations between some metabolic and sensory traits, such as citric acid with tomato smell, glycine with tomato smell, and granulosity with dry matter, suggests a basis for more targeted investigations in the future.

Key words: Flavour, metabolic profiling, network analysis, sensory analysis, tomato.

Introduction

Flavour is a very complex trait that is affected by many genetic components and non-genetic factors, not all of which are known or well understood (Baldwin *et al.*, 2000; Tandon *et al.*, 2003; Goff and Klee, 2006). A complex mixture of sugars, acids, amino acids, minerals, and volatile compounds contributes to the characteristic flavour of fresh tomato fruits (Stevens *et al.*, 1977; Petrò-Turza, 1987; Baldwin *et al.*, 1991*a*, *b*, Buttery, 1993). The concentrations of these molecules may significantly affect flavour acceptability (Malundo *et al.*, 1995).

Recent scientific discoveries regarding tomato fruit flavour components (Saliba-Colombani *et al.*, 2001; Fulton *et al.*, 2002; Chaib *et al.*, 2007) have encouraged efforts to improve this trait genetically. Quantitative trait locus (QTL) studies in tomato have found that that few chromosome regions control the variation of sensory and biochemical traits (Causse *et al.*, 2002). Tieman *et al.* (2006) identified a number of QTLs that reproducibly alter the composition of volatiles and chemicals that contribute to overall fruit flavour, whilst Chaib *et al.* (2007) identified several tomato fruit parameters associated with sensory texture attributes useful to improve knowledge of their genetic control. Several recent studies have developed mathematical relationships between sensory descriptors and instrumental measures of fresh tomato flavour (Baldwin *et al.*, 1998; Krumbein and Auerswald, 1998; Auerswald *et al.*, 1999; Gajc-Wolska *et al.*, 2000; Krumbein *et al.*, 2000; Maul *et al.*, 2000; Tandon *et al.*, 2003) to provide reliable analytical tools.

Technological developments have considerably extended our ability to describe complex biological systems;

* To whom correspondence should be addressed. E-mail: ercolano@unina.it

high-throughput methods now allow the simultaneous analysis of several metabolites. The development of high-throughput data collection techniques helps to determine how and when these molecules interact with each other. In the global growth of 'omic' strategies in plants, high-throughput metabolite screening techniques will generate large volumes of analytical data that can be added to the rapidly expanding collections of gene sequence, phenotypic, and gene expression data. The use of global data rather than single trait analysis can be an effective way to visualize complex phenomena in a single experiment. After all, biological functions can rarely be attributed to an individual molecule.

Network analysis has proved to be a powerful tool to distil data into meaningful information. Biological networks give a visual representation of biological systems, capturing their essential characteristics and interactions. Many systems of current interest to the scientific community can be usefully represented as networks (Kauffman, 1969; Barabási and Oltvai, 2004; Schauer *et al.*, 2006). In the network, the traits are represented by nodes that are connected by links, with each link representing the interactions between two components.

In this work, tomato metabolite profiling was performed on eight different genotypes in parallel with plant phenotype characterization and fruit sensory analysis in order to investigate the simultaneous expression of fruit traits. To identify relationships among tomato metabolites, sensory profile analysis, and agronomic features, a biological network was constructed. Various types of interactions, including amino acid networks, phenotypic and metabolic associations with sensory attributes, and links among metabolites of organoleptic importance, were revealed and essential relationships visualized. In this way, it was possible to reduce data complexity by focusing on key information of the full data set.

Materials and methods

Plant material and growth

The eight indeterminate tomato genotypes utilized in this work were as follows: six traditional tomato landraces (100 Sch, Ves 2001, Sor Art, Sor Adg, Sm Sch, and Sm Sel 6), one fresh market variety (MONEY MAKER), and one processing variety (E6203). These genotypes were grown in randomized, replicated plots in two different sites in southern Italy (Sorrento and Sarno) during the summer of 2005. Young seedlings (~1 month old) were planted at the end of April in a randomized complete block design with two replications. Plants were grown under the standard tomato field procedures used for the area. Ripe fruits from all plants for each line were harvested three times, and fruit yield (g per plant), number of fruits, and morphological traits (fruit polar and equatorial diameters) recorded for single plants. At the three different harvesting times, a sample for each replicate (10 plants) of 2–6 kg was obtained by pooling fruits belonging to each genotype. Random pieces of fruits were used to conduct sensory evaluation. Furthermore, the fruits were homogenized, divided into aliquots, and stored at –20 °C to determine chemical and biochemical parameters.

Chemicals

All solvents used for HPLC analysis were purchased from Merck (Darmstadt, Germany). The malic and fumaric acid standards were from ICN Biomedical Inc., ascorbic acid and citric acid were from Sigma (CA, St Louis, MO, USA), and the amino acids were supplied by Bachem (Switzerland).

Physical and chemical analysis

In order to perform physical, chemical, and biochemical analyses, a homogenized mix of fruits of the three field harvests of each genotype was obtained. The following parameters were determined on all samples in duplicate: pH at 20 °C (HI 9017 Microprocessor pHmeter, Hanna Instruments), colour (L, a, b), refractive index at 20 °C (°Brix), total solids, total acidity, chloride ions, ash, organic acids, and amino acids.

The colour parameters 'a', (green-to-red coordinate) 'b' (blue-to-yellow coordinate), and 'L' (lightness) were determined for the various samples with a Hunter Lab D25 A Optical Sensor-Reston (Virginia, USA). The soluble solid concentration in the fruit was estimated by evaluating the degree of Brix, which was determined on the homogenate by an RFM330 Refractometer (Bellingham Stanley Ltd, UK). Total solids (dry matter content) were estimated by drying 5 g of fresh fruit in an oven (Ehret) set at 70 °C until constant weight was reached. Results were expressed as percentages of fresh weight. Total acidity and chloride ions were analysed with a Crison TT2050 pH-meter. Ash content was calculated from the weight of the sample after burning at a temperature of 105 °C overnight (Clarke and Walker, 1975).

Organic acids

The organic acids (malic, citric, ascorbic, and fumaric) were determined by HPLC analysis. Briefly 0.1 g of lyophilized sample was added to 5 ml of 0.008 N H_2SO_4/H_2O, agitated for 1 min, and centrifuged at 4000 rpm for 5 min at 4 °C. Aliquots of 2 ml of the supernatant were collected and centrifuged at 12 000 rpm for 2 min at 4 °C. An aliquot of the extract was used for analysis by HPLC configured with LC-10AD pumps, an SLC10A system control, a diode array UV-VIS detector (Shimadzu, Japan), and a Synergy Hydro column (4 μm, 250 mm×4.6 mm; Phenomenex). The organic acids were eluted with 0.008 N H_2SO_4/H_2O at 1.0 ml min^{-1} under isocratic conditions at 210 nm for malic, citric, and fumaric acids, and at 245 nm for ascorbic acid. Extraction was repeated twice for each sample. The data obtained were expressed as milligrams of organic acids per 100 g of fresh matter.

Amino acids

In order to evaluate the amino acid content, 25 g of freeze-dried tomato samples were dissolved in 15 ml of deionized water and centrifuged at 4000 rpm for 15 min. The supernatant was filtered and centrifuged using a Centricon YM-3 (Millipore, USA). A 500 µl aliquot of filtrated sample was dried and dissolved in 500 µl of borate buffer (0.1 M, pH 10.4). The solution was mixed with Fmoc reagent (500 µl, 5.8 mM in acetone) (Gartenmann and Kochlar, 1999). The mixture was extracted twice with 2 ml of hexane/ethyl acetate (80:20). The aqueous phase containing the Fmoc derivatives was analysed by RP-HPLC interfaced with an ESI-MS (electrospray ionization–mass spectrometer; API-100 Sciex, Canada), using the following conditions for HPLC and MS.

HPLC: Liquid chromatography (LC) analyses were performed using two series 200 micro pumps (Perkin Elmer; Canada). A 250×4.6 mm Luna 5 µm C_{18} column (Phenomenex, USA) was used. Eluents were water+0.05% trifluoroacetic acid (TFA; solvent A) and acetonitrile+0.05% TFA (solvent B). The Fmoc derivatives were separated using the following linear gradient: 30–50% B in 15 min, 50–100% B in 20 min, and 5 min isocratic elution at 100% B. The LC flow rate was set at 0.8 ml min^{-1} and after the split 50 µl min^{-1} was sent to the mass spectrometer. The injection volume was 50 µl.

MS: The analyses were performed using an API 100 single-quadruple mass spectrometer equipped with an ESI source in positive mode. The operating parameters were as follows: capillary voltage 5000 V, orifice voltage 100 V. Acquisition was performed in SIM (single ion monitoring) using a dwell time of 300 ms.

Sensory analysis

Sensory analyses were performed by a trained panel of 10 judges. The panel worked in a sensory laboratory under defined (temperature and light) conditions in single booths with computer equipment.

For each genotype, 10 different attributes were revealed: one related to appearance (redness), one to smell (tomato smell), three to taste (sweetness, saltiness, and sourness), one to flavour (tomato flavour), and four to texture (hardness, juiciness, granulosity, and skin resistance). Determination of the intensity of sensory perception by the trained panel was carried out twice for each type of product with the use of unstructured line scales with the anchor points 0—not perceptible, and 100—strongly perceptible.

Statistical analysis

Statistical analysis was performed using the R statistical software (R Development Core Team, 2008). Statistical analysis was divided into two steps: first the presence of significant variation among varieties (genotypes) of the 37 variables was verified, and then insight was gained into their possible inter-relationships through visualization of a network structure. A complex descriptive technique, previously used in the literature, which gives an immediate graphical display of the underlying complex relationships was employed (Ursem *et al.*, 2008). Other multivariate techniques (principal component analysis and multidimensional scaling) were also implemented. The results were not so easily interpretable, although in some way they confirmed some of the findings reported with the social network, and it was therefore decided not to report them.

In order to ascertain the effect of genotypes and location, a two-factor analysis of variance (ANOVA) model was applied using a significance level of $P < 0.01$ and $P < 0.001$. In order to reduce the effect of confounding factors, such as location and genotype, variables with a significant effect of $P < 0.001$ were discarded for subsequent analysis.

Pearson's correlations between all trait pairs were calculated and the significance of their associations was tested with a *t*-test at a significance level of 0.05. Relying on a correlation matrix, a social network (Wasserman and Faust, 1994) was set up using scripts kindly provided by Dr Dani Zamir (Hebrew University of Jerusalem). The network was then constructed from the correlation matrix resulting from 35 variables measured in eight tomato genotypes harvested in two locations, considered as a homogenous sample (a sample for which we can assume that each measurements is a random sample from a unique distribution).

In the network structure, vertices correspond to a trait and links between two vertices correspond to significant correlations between these two traits. Two nodes were connected by a link if the correlation (positive or negative) between the components was significant at level $\alpha=0.05$. This network was then used as input for a cartographic algorithm (Guimerà and Nunes Amaral, 2005), which allows the network to be divided into modules or groups of vertices that are more connected between themselves than to nodes of other modules. To test the robustness of the method, the procedure was repeated 100 times using different starting points. The modularity of the network, defined as a quantity which becomes larger with an increase in the number of edges in a cluster and with a decrease in the number of edges between two different clusters, has been computed according to the following formula:

$$\sum_{s=1}^{Nm}\left[\frac{l_s}{L}-\left(\frac{d_s}{2K}\right)^2\right]$$

where Nm is the number of modules in the network, L the number of internal and external connections in the network, d_s the number of external connections of module s, and l_s the number of internal connections between modules.

Results

To assess relationships which exist among characteristics involved in tomato fruit organoleptic quality, 37 agronomic,

biochemical, and sensory parameters were measured in eight different genotypes (six traditional tomato landraces, one fresh market variety, and one processing variety) harvested in two different locations. As initial data exploration, the extent of trait variation among genotypes was analysed by ANOVA (Supplementary Table S1 available at *JXB* online). The effects of both the genotype and the location on each variable were investigated.

For each genotype, Table 1 reports the average value and the standard deviation of all traits measured in the two fields. Significant variations among varieties were found for fruit polar diameter (P <0.001), dry matter (P <0.01), aspartate content (P <0.001), and ascorbic acid content (P <0.01), while between fields significant differences were found for asparagine and fumaric acid content (P <0.01). With regards to the sensory attributes, the only significant differences among genotypes detected by the panel were those related to juiciness (P <0.01) and skin resistance (P <0.01) for genotypes, and to flavour (P <0.01) for location.

In order to perform network analysis, the variables (aspartate and polar diameter) for which large significant effects (P <0.001) of genotypes and location were detected by ANOVA were discarded. It was thus possible to reduce the effects of confounding factors that could produce bias in network data elaboration.

Table 1. Evaluation of agronomic, physicochemical, biochemical, and sensory traits of tomato fruits in eight tomato varieties harvested in two different fields

Traits that are significantly different (P <0.001) are in bold.

Item	Trait	Genotypes							
		SM Sch	**Ves 2001**	**Sel 6**	**Sor Adg**	**M. M.**	**E6203**	**Sor Art**	**100 Sch**
Agronomic traits	Total yield (g per plant)	1154±589[a]	2080±1717	773±635	351±216	1830±991	1004±153	596±71.2	1903±761
	Fruits (*n*)	37.7±26.1	142±44.6	17.8±10.2	2.90±2.96	42.6±22.0	20.0±2.26	4.50±0.92	85.1±22.7
	Equatorial diameter (cm)	3.66±0.00	3.27±0.24	3.49±0.18	7.53±0.51	4.80±0.45	4.70±0.14	4.17±4.99	3.40±0.11
	Polar diameter (cm)	**6.06±0.34**	**4.26±0.14**	**7.37±1.79**	**7.10±0.09**	**4.53±0.41**	**5.14±0.68**	**6.79±0.65**	**5.03±0.01**
Physicochemical traits	Colour 'L'	30.4±6.47	41.5±8.55	36.7±5.67	39.1±10.8	50.5±1.37	33.1±1.57	39.8±4.76	41.5±2.19
	Colour 'a'	28.4±3.56	26.7±2.37	28.5±2.40	23.5±8.75	18.1±3.60	28.7±4.63	25.4±8.55	24.4±2.57
	Colour 'b'	9.71±1.86	13.5±0.84	10.9±1.06	7.92±3.47	11.4±2.79	11.7±2.88	11.6±0.62	12.7±3.55
	pH	4.21±0.22	4.07±0.07	4.18±0.16	4.54±0.08	4.36±0.34	4.10±0.22	4.07±0.09	4.14±0.10
	Acidity (%)	0.40±0.02	0.44±0.09	0.45±0.15	0.33±0.00	0.45±0.15	0.48±0.15	0.37±0.30	0.36±0.06
	Chloride ions (%)	0.04±0.01	0.04±0.00	0.03±0.01	0.02±0.02	0.03±0.01	0.07±0.02	0.03±0.03	0.03±0.00
	°Brix	5.05±0.77	5.65±0.91	5.60±1.69	4.55±0.21	4.55±0.21	4.95±1.20	4.80±0.00	4.80±0.70
	Dry matter (%)	5.79±0.58	8.34±0.55	6.15±0.88	3.70±0.98	4.45±0.07	5.69±1.54	6.12±0.34	6.34±1.76
	Ash (%)	0.49±0.05	0.46±0.05	0.45±0.10	0.45±0.19	0.45±0.05	0.56±0.27	0.46±0.02	0.44±0.12
Biochemical traits	Malic acid (mg 100g^{-1} FW[b])	71.8±44.6	127±43.6	71.0±2.82	67.4±27.5	122±14.2	78.5±53.9	76.0±8.62	124±11.2
	Citric acid (mg 100 g^{-1} FW)	315±50.6	337±85.0	424±139	198±77.1	209±12.4	286±154	396±132	279±63.0
	Ascorbic acid (mg 100 g^{-1} FW)	5.16±3.04	7.00±0.05	0.04±0.06	0.00±0.00	1.47±2.08	1.83±2.60	0.00±0.00	1.22±1.72
	Fumaric acid (mg 100 g^{-1} FW)	0.05±0.07	0.27±0.15	0.20±0.14	0.38±0.38	0.24±0.10	0.20±0.28	0.36±0.34	0.27±0.12
	His (mg 100 g^{-1} FW)	20,6±12,0	17.4±7.91	25.3±9.61	21.6±10.6	16.3±7.84	18.6±14.1	10.1±3.81	9.85±6.29
	Lys (mg 100 g^{-1} FW)	31.5±10.8	26.1±1.97	33.6±2.82	29.6±5.44	16.3±4.59	28.3±3.11	20.6±17.8	20.5±11.6
	Arg (mg 100 g^{-1} FW)	53.2±27.2	48.3±3.18	48.5±0.98	68.1±46.8	28.2±7.49	51.4±1.97	27.1±23.7	32.1±14.4
	Gln (mg 100 g^{-1} FW)	208±82.1	126±60.5	219±28.6	175±90.2	109±18.2	232±89.1	183±112	164±93.2
	Ser (mg 100 g^{-1} FW)	54.7±13.2	47.7±1.76	58.6±13.5	47.3±11.1	34.1±9.97	49.5±3.53	44.3±32.6	42.3±29.6
	Asp (mg 100 g^{-1} FW)	**122±39.9**	**118±78.9**	**106±5.51**	**137±105**	**52.3±22.0**	**115±74.5**	**144±92.4**	**80.4±36.4**
	Glu (mg 100 g^{-1} FW)	62.8±82.5	140±67.4	76.9±15.8	93.4±19.7	69.2±0.98	127±85.6	88.6±1.62	63.4±52.7
	Asp (mg 100 g^{-1} FW)	24.2±3.74	20.5±8.41	18.3±5.79	22.1±0.84	8.75±3.74	27.2±0.98	23.5±0.77	121±29.1
	Thr (mg 100 g^{-1} FW)	34.6±2.26	45.7±3.11	36.7±4.38	34.2±10.4	23.2±4.80	41.9±2.75	20.8±19.9	31.6±18.1
	Gly (mg 100 g^{-1} FW)	21.5±0.91	20.8±12.5	18.4±0.16	17.7±2.82	9.45±13.3	10.1±1.06	10.1±14.3	20.4±0.98
Sensory traits	Redness	45.0±26.8	61.5±3.53	53.5±4.94	43.5±7.77	68.0±7.07	60.5±0.70	47.5±27.5	53.5±7.77
	Tomato smell	40.0±4.24	44.5±3.53	42.5±0.70	40.0±2.82	51.0±5.65	40.5±0.70	37.0±7.07	47.0±1.41
	Sweetnees	25.5±2.12	37.5±6.36	31.0±2.82	28.5±2.12	28.0±8.48	23.0±8.48	33.5±4.50	35.0±4.50
	Saltiness	23.0±1.41	24.0±0.00	25.0±4.24	27.0±2.82	27.5±2.12	28.5±3.53	22.5±4.95	24.0±2.82
	Sourness	27.5±3.53	20.5±0.70	30.0±2.82	32.5±6.36	29.5±9.19	29.0±4.25	21.5±3.54	29.5±2.12
	Flavour	35.5±2.12	41.0±5.65	41.0±7.07	38.5±10.6	40.0±9.89	38.5±2.12	38.0±5.65	36.5±4.94
	Hardness	32.5±9.19	30.0±2.82	38.0±1.41	34.0±1.41	36.5±6.36	32.0±7.07	29.0±2.82	34.5±3.53
	Juiciness	35.0±4.24	46.0±4.24	40.5±2.12	42.0±0.00	46.5±24.7	45.0±11.3	39.5±14.8	44.5±3.53
	Granulosity	35.5±2.12	30.0±5.65	39.5±5.65	35.5±14.8	27.0±1.41	29.0±5.65	25.5±0.70	30.0±2.82
	Skin resistance	49.0±2.82	55.0±2.82	51.5±4.94	39.5±0.70	58.5±0.70	52.0±5.65	33.5±4.94	51.5±0.70

[a] Values are presented as mean ±SD for two different locations; single field values are derived from two independent determinations.
[b] Fresh weight.

Single homogeneous samples were built up using all measurements available for the remaining 35 variables and employed for network analysis. Figure 1 showed the variation identified in homogeneous samples obtained for sensory attributes.

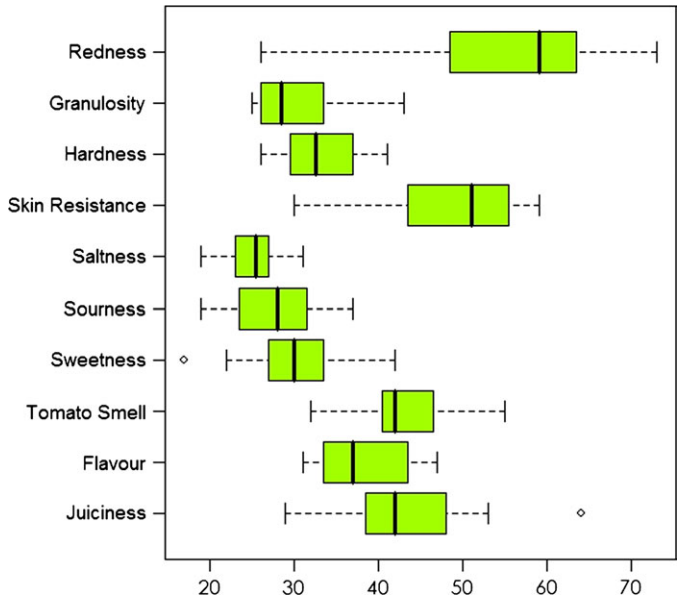

Fig. 1. Box plots of the sensory data, showing variation within a homogeneous sample.

Correlation analysis among the 35 variables resulted in a matrix containing 1225 correlations with a minimum of –0.857 and a maximum of 0.924. On the basis of this matrix, a biological network was constructed where each trait is represented as a node possibly connected to any other node. The network consists of 35 nodes and 74 links corresponding to the 74 significant (positive or negative) correlations. Two modules with a large number of internal links, two with few connections inside and outside the module, and one module with only three sensory characteristics were observed (Fig. 2). In order to interpret the meaning of the information from the biological network, knowledge of the role of each node is of crucial importance. The network displayed three nodes interconnected between themselves (dry matter, pH, and °Brix) and with other traits, and nodes with widely different degrees (scales) mainly inside the same module.

The first module comprised most amino acids and colour components 'L' and 'a'. This module has a higher value of specific modularity (0.196). Indeed, the traits belonging to this module were highly interconnected, while they showed few links (only four) outside the module. As expected, a strong negative correlation was obtained between the two colour components, 'a' and 'L'.

The second module included the amino acids glutamate and asparagine, citric acid (the major acid present in tomato fruits) and fumaric acid, physicochemical traits (acidity, ash, chloride ions, and °Brix), and three sensory attributes: tomato smell, juiciness, and overall aroma. This group has

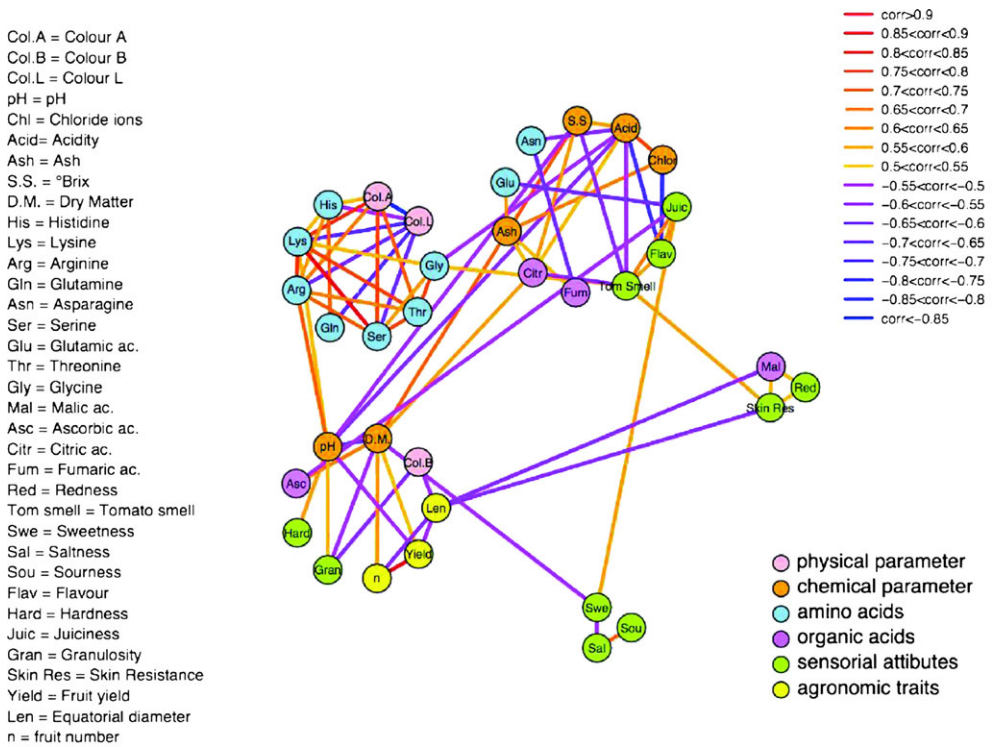

Fig. 2. Map of the combined agronomic, metabolic, and sensory tomato trait network. Each trait (node) is represented by a circle and coloured as follows: pink, physical parameter; orange, chemical parameter; sky blue, amino acids; mauve, organic acids; green-yellow, sensory attributes; and yellow, agronomic traits. Interactions are indicated with lines: red represents positive correlations; blue represents negative correlations.

the second highest value of modularity (0.155). The strongest negative correlation was found between flavour and both acidity and chloride ion content, while the strongest positive correlation was shown between acidity and chloride ions. Tomato smell was affected by °Brix, acidity, and citric acid, while juiciness was positively correlated with glutamate content. Interestingly, the three sensory parameters that were grouped in this module are interconnected, while the other biochemical traits chiefly showed links with nodes belonging to other modules.

The third module has a lower modularity (0.116) since the nodes showed few connections among internal and external traits. The traits belonging to this module were quite heterogeneous: the three agronomic traits (fruit yield, equatorial diameter, and fruit number per plant), ascorbic acid, the 'b' colour component, pH, and dry matter, and parameters evaluated through the panel test such as hardness and granulosity. Dry matter showed the highest number of connections with the other nodes inside (five) and outside (three) the module. The highest positive correlation was found between fruit yield (g per plant) and fruit number per plant. The agronomic traits did not seem to influence sensory parameters directly, with the exception of skin resistance.

The last two modules each grouped only three characteristics. The first included the high interconnecting traits: malic acid, skin resistance, and redness, showing a modularity of 0.036. Skin resistance showed two links with equatorial diameter and tomato smell outside the module. The second included only three sensory attributes: sweetness, saltiness, and sourness. It has the lowest specific modularity (0.025), since it showed only two links inside the module and two outside the module. Saltiness is connected with both sweetness and sourness. No major connections of these sensory attributes with other traits were identified, except for sweetness with juiciness and dry matter.

Discussion

Assessing traits that contribute to define the target flavour and designing strategies to improve it is a long-term objective of tomato breeding programmes. In order to identify key relationships among tomato metabolites, sensory profile analysis, and agronomic features, a biological network was constructed. The graphic representation of this network revealed various types of interactions useful to visualize essential relationships among fruit traits.

As previously reported, amino acids were strongly intercorrelated: in the network they all (with the exception of glutamate and asparagine) belong to the same module and show many intercorrelations. The high interconnectivity of the amino acids comes as no surprise given the exquisite multilevel regulation mechanisms operating on their metabolism. This result agreed with several studies which reported that the network of amino acid metabolism is subject to a high degree of metabolic regulation (Galili, 1995; Galili and Hofgen, 2002). For instance, lysine and threonine, which showed a high level of correlation, were both synthesized in plants from aspartate by two different pathways.

This first module also included the two colour components 'L' and 'a', which showed a very strong negative correlation (less than –0.85). In line with several other studies, the 'a' value showed a linear correlation with the ripening stages of the tomatoes. The lightness factor 'L', on the other hand, decreased during the first five ripening stages and then remained constant (Arias *et al.*, 2000; Raffo *et al.*, 2002). Given that amino acids clustered with colour components 'a' and 'L', similar co-regulation of amino acid synthesis during the ripening stages could also be hypothesized. During the whole ripening process, modification of amino acid metabolism was also observed (Faurobert *et al.*, 2007). Changes in amino acid composition have been reported, influenced by the enzyme degradation process related to tomato fruit shelf-life (Boggio *et al.*, 2000; Pratta *et al.*, 2004).

Interestingly, the amino acid glutamate was included in another module with five biochemical traits and three sensory attributes. It is one of the predominant amino acids found in tomato and can have an effect on overall tomato taste (Fuke and Shimizu, 1993). Petrò-Turza and Teleky-Vamossy (1989) showed that the addition of the amino acids glutamate and aspartate to a model juice made of mineral salts, sugars, and acids leads to a significant improvement in taste characteristics. This finding should be taken into account in balancing tomato fruit characteristics as amino acids are precursors of important flavour volatiles (Tieman *et al.*, 2006).

Looking at the other biochemical traits, dry matter showed links with ascorbic acid and pH within the same module and with °Brix and citric acid outside the module. This evidence is not surprising as ~60% of tomato dry matter consists of sugars and organic acids. The °Brix was also linked positively to both citric acid and acidity.

Interestingly, the sensory attributes were distributed in four modules, but none showed a high number of links. The construction of this network suggests that while there are clearly interconnections among specific sensory traits, few strong relationships between sensory perception and specific biochemical traits can be identified. In particular, sourness and saltiness failed to correlate with any biochemical traits, while sweetness showed connectivity only with dry matter and ascorbic acid, but not with °Brix.

Previous studies reporting correlations between sensory and physicochemical and biochemical traits indicated that these were loose except for some expected correlations such as sweetness with sugar content (Baldwin *et al.*, 1998; Causse *et al.*, 2002), and sourness with titratable acidity (Stevens *et al.*, 1977, 1979; Causse *et al.*, 2004) and pH (Tandon *et al.*, 2003). A high degree of connectivity between the mentioned traits and sensory attributes was not found in the present study. It could be hypothesized that regulatory factors, responsible for balancing several classes of metabolites, act on different circuits determining the perception of tomato flavour. Besides, it is known that some compounds have a threshold effect which might be difficult to detect and that organoleptic perception results from the overall interaction among fruit components in the mouth.

The present data suggest that to gain a more comprehensive understanding of tomato flavour components it is important to assess functional interactions in the fruit as a whole. For instance, agronomic traits evaluated in this study did not show many connections, but tight positive connections between total yield and fruit number with dry matter were displayed. Schauer *et al.* (2006) reported that plants with a lower harvest index have a high °Brix value. They suggest that large numbers of metabolite traits are controlled by source–sink partitioning. This implies that targeted manipulation of the content of metabolites of central pathways can also be influenced by the plant translocation process.

The main feature of the network of tomato characteristics was the presence of three nodes interconnected between themselves (dry matter, pH, and °Brix) and with other traits, and nodes with widely different degrees (scales) mainly inside the same module. This finding confirms that such traits represent key parameters for tomato fruit quality, and their variance can influence variance in other traits. Given the number of connections that each node establishes with other traits, it should be pointed out that dry matter has both a high total number of connections (eight) inside the module (five) and a high number of connections (three) to nodes external to the module to which it belongs. In contrast, all other nodes, even those that are highly connected, are predominantly linked to other nodes within the same module.

In-depth interpretation and understanding of the network connections is not a trivial task, as many of the interactions and many network features were previously unknown. However, dry matter, pH, and °Brix were identified as important drivers of fruit tomato quality components. Moreover, few metabolic traits seem to have a direct influence on important sensory traits, such as citric acid on tomato smell, glutamic acid on juiciness, and dry matter on granulosity. Use of mutants for these traits may lead to the identification of metabolic pattern changes that alter tomato fruit composition and the perception of sensory attributes. Preliminary experiments conducted on tomato mutants for citric acid showed 44 significant metabolite variations (out of 53 assessed) between normal fruits and mutants lacking citric acid (data not shown).

Sensory attributes contributing to organoleptic perception, such as sweetness, saltiness, and sourness for taste, and hardness and granulosity for texture, are grouped together in the different modules, with weak connectivity among themselves and with other traits. This finding was somewhat unexpected and should also be taken into account.

In conclusion, the modular partitioning of tomato characteristics involved in fruit organoleptic perception captured essential fruit parameters that interact with different class traits. A number of interesting links were identified. These links can be both positive and negative contributors to tomato flavour and can have direct implications for crop improvement strategies. This could be relevant when developing new tomato varieties to be launched on the market for their nutritional and organoleptic characteristics (Frusciante *et al.*, 2007). Further research is required to clarify the biological significance of such observations. Only with a larger sample size will it be possible to challenge the strength of the present findings and improve the understanding of complex interactions with inferential methods relying on a comprehensive statistical model.

Supplementary data

Supplementary Table S1 reports the variance value, the *F*- and *P*-values of agronomic, physicochemical, biochemical, and sensory traits investigated, the minimum and maximum values, the average value, and the standard deviation of all genotypes considered as a homogeneous sample.

Acknowledgements

The authors wish to thank Dr Alisdair Fernie, Dr Anna Lytovchenko, and Dr Amalia Barone for their useful suggestions; Mark Walters for editing the manuscript; and Michele De Martino for his technical assistance. This work was performed within the framework of the Project 'Risorse genetiche di organismi utili per il miglioramento di specie di interesse agrario e per un'agricoltura sostenibile' funded by the Ministry for Agricultural and Forestry Policy (MiPAF) Contribution no. 194 from the DISSPAPA.

References

Arias R, Lee TC, Logendra L, Janes H. 2000. Correlation of lycopene measured by HPLC with the L*, a*, b* colour readings of a hydroponic tomato and the relationship of maturity with colour and lycopene content. *Journal of Agricultural Food Chemistry* **48,** 1697–1702.

Auerswald H, Peters P, Brückner B, Krumbein A, Kuchenbuch R. 1999. Sensory analysis and instrumental measurement of short-term stored tomatoes (*Lycopersicon esculentum* Mill.). *Postharvest Biology Technology* **15,** 323–334.

Baldwin EA, Nisperos-Carriedo MO, Baker R, Scott JW. 1991*a*. Qualitative analysis of flavour parameters in six Florida tomato cultivars. *Journal of Agricultural Food Chemistry* **39,** 1135–1140.

Baldwin EA, Nisperos-Carriedo MO, Moshonas MG. 1991*b*. Quantitative analysis of flavour and other volatiles and for certain constituents of two tomato cultivars during ripening. *Journal of the American Society for Horticultural Science* **116,** 265–269.

Baldwin EA, Scott JW, Einstein MA, Malundo TMM, Carr BT, Shewfelt RL, Tandon KS. 1998. Relationship between sensory and instrumental analysis for tomato flavor. *Journal of the American Society for Horticultural Science* **125,** 906–915.

Baldwin EA, Scott JW, Shewmaker CK, Schuch CK. 2000. Flavour trivia and tomato aroma: biochemistry and possible mechanism for control of important aroma components. *Horticultural Science* **35,** 1013–1022.

Barabási AL, Oltvai ZN. 2004. Network biology: understanding the cell's functional organization. *Nature Reviews Genetics* **5,** 101–113.

Boggio SB, Palatnik JF, Heldt HW, Valle EM. 2000. Changes in the amino acid composition and nitrogen metabolizing enzymes in

ripening fruits of Lycopersicon esculentum Mill. *Plant Science* **159,** 125–133.

Buttery RG. 1993. Quantitative and sensory aspects of flavor of tomato and other vegetables and fruits. In: Acree TE, Teranishi R, eds. *Flavor science: sensible principles and techniques*. Washington, DC: American Chemical Society, 259–286.

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.

Chaib 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,** 1–11.

Clarke RJ, Walker LJ. 1975. The interrelationships of potassium contents of green, roasted and instant coffees. In: *Proceedings of the Seventh International Scientific Colloquium on Coffee*. Hamburg: ASIC-Association Scientifique Internationale du Café, 159–163.

Faurobert M, Mihr C, Bertin N, Pawlowski T, Negroni L, Sommerer N, Causse M. 2007. Major proteome variations associated with cherry tomato pericarp development and ripening. *Plant Physiology* **143,** 1327–1346.

Frusciante L, Carli P, Ercolano MR, Pernice R, Di Matteo A, Fogliano V, Pellegrini N. 2007. Antioxidant nutritional quality of tomato. *Molecular, Nutritional and Food Research* **51,** 609–617.

Fuke S, Shimizu T. 1993. Sensory and preference aspects of umami. *Trends in Food Science and Technology* **4,** 246–251.

Fulton TM, Bucheli P, Voirol E, Lopez J, Petiard V, Tanksley SD. 2002. Quantitative trait loci (QTL) affecting sugars, organic acids and other biochemical properties possibly contributing to flavour, identified in four advanced backcross populations of tomato. *Euphytica* **127,** 163–177.

Gajc-Wolska J, Skqpski H, Szymczak JA. 2000. Chemical and sensory characteristics of the fruits of eight cultivars of field grown tomato. *Acta Physiologiae Plantarum* **22,** 369–373.

Galili G. 1995. Regulation of lysine and threonine synthesis. *The Plant Cell* **7,** 899–906.

Galili G, Hofgen R. 2002. Metabolic engineering of amino acids and storage proteins in plants. *Metabolic Engineering* **4,** 3–11.

Gartenmann K, Kochhar S. 1999. Short-chain peptide analysis by high-performance liquid chromatography coupled to electrospray ionization mass spectrometer after derivatization with 9-fluorenylmethyl chloroformate. *Journal of Agricultural Food Chemistry* **47,** 5068–5071.

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

Guimerà R, Nunes Amaral LA. 2005. Functional cartography of complex metabolic networks. *Nature* **433,** 895–900.

Kauffman SA. 1969. Metabolic stability and epigenesis in randomly constructed genetic nets. *Journal of Theoretical Biology* **22,** 437–467.

Krumbein A, Auerswald H. 1998. Characterization of aroma volatiles in tomatoes by sensory analyses. *Nahrung/Food* **42,** 395–399.

Krumbein A, Peters P, Widell S, Auerswald H, Brückner B, Linke M. 2000. Sensory attributes and flavour compounds of short-term stored vegetables. In: Florkowski WJ, Prussia SE, Shewfelt RL, eds. *Fruit and vegetable quality: an integrated view*. Lancaster, UK: Technomic Publishing Inc, 241–249.

Malundo TMM, Shewfelt RL, Scott JW. 1995. Flavour quality of fresh tomato (*Lycopersicon esculentum* Mill.) as affected by sugar and acid levels. *Postharvest Biology and Technology* **6,** 103–110.

Maul F, Sargent SA, Sims CA, Baldwin EA, Balaban MO, Huber DJ. 2000. Tomato flavour and aroma quality as affected by storage temperature. *Journal of Food Science* **65,** 1228–1237.

Pratta G, Zorzoli R, Boggio SB, Picardi LA, Valle EM. 2004. Glutamine and glutamate levels and related metabolizing enzymes in tomato fruits with different shelf-life. *Scientia Horticulturae* **100,** 341–347.

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

Petrò-Turza M, Teleky-Vamossy G. 1989. Study on taste substances of tomato. Part 3. Sensory evaluations. *Die Nahrung* **33,** 387–394.

R Development Core Team. 2008. *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. ISBN 3-900051-07-0. http://www.R-project.org.

Raffo A, Leopardi C, Fogliano V, Ambrosino P, Salucci M, Gennaro L, Bugianesi R, Giufridda F, Quaglia G. 2002. Nutritional value of cherry tomatoes (*Lycopersicon esculentum* cv. Naomi F1) harvested at different ripening stages. *Journal of Agricultural Food Chemistry* **50,** 6550–6556.

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

Stevens MA, Kader AA, Albright-Holton M. 1979. Potential for increasing tomato flavor via increased sugar and acid content. *Journal of the American Society for Horticultural Science* **104,** 40–42.

Stevens MA, Kader AA, Albright Holton M, Algazi M. 1977. Genotypic variation for flavor and composition in fresh market tomatoes. *Journal of the American Society for Horticultural Science* **102,** 680–689.

Tandon KS, Baldwin EA, Scott JW, Shewfelt RL. 2003. Linking sensory descriptors to volatile and nonvolatile components of fresh tomato flavor. *Journal of Food Science* **68,** 2366–2371.

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

Ursem R, Yury T, Bovy A, van Berloo R, van Eeuwijk F. 2008. A correlation network approach to metabolic data analysis for tomato fruits. *Euphytica* **161,** 181–193.

Wasserman S, Faust K. 1994. *Social network analysis*. Cambridge: Cambridge University Press.