Breeding tomato flavor: modeling consumer preferences of tomato landraces.

Villena, J.[a], Moreno, C.[a], Roselló, S.[b], Beltran, J.[c], Cebolla-Cornejo, J.[d], Moreno, M.M.[a*]

[a]*University of Castilla-La Mancha, Higher Technical School of Agricultural Engineering in Ciudad Real, Ronda de Calatrava 7, 13071, Ciudad Real, Spain*

[b]*Joint Research Unit UJI-UPV - Improvement of agri-food quality. Agricultural Sciences and Natural Environment Department, Universitat Jaume I, Avda. Sos Baynat s/n, 12071 Castelló de la Plana, Spain*

[c]*Research Institute for Pesticides and Water (IUPA). Universitat Jaume I, Avda. Sos Baynat s/n, 12071 Castelló de la Plana, Spain*

[d]*Joint Research Unit UJI-UPV - Improvement of agri-food quality. COMAV. Universitat Politècnica de València, Cno. de Vera s/n, 46022 València, Spain*

*Corresponding author

Abstract

Tomato landraces are highly appreciated by consumers, who are willing to pay a price premium for them. But little is known regarding sensory perception and its relationship with fruit composition. The Spanish variety "Moruno" was selected as a model for this purpose. A collection of 30 populations was grown in different environments and evaluated by a consumer panel. Partial least square (PLS) models were then developed relating determinant flavor descriptors (sweetness, sourness, taste and aroma intensity, aftertaste persistence and agreeability, and overall flavor acceptability) with compositional variables such as soluble solids, pH, titratable acidity, individual sugars, organic acids, volatiles, and derived variables. PLS models identified relationships that had not been uncovered with correlation and simple regression analysis and offered low cross-validation errors (<15% of the range of variation). Although plant yield negatively affected sensory perception, it was possible to identify populations with a good combination of both traits.

Highlights

Volatiles and soluble solids interact in all the sensory descriptors

PLS models uncover the role of compounds not highlighted in correlation analysis

Model error (RMSECV) represented less than 15% of the range of variation

Chemical compounds studied in this article

Fructose (PubChem CID: 5984); Glucose (PubChem CID: 5793); Citric acid (PubChem CID: 311); Malic acid (PubChem CID: 525); Glutamic acid (PubChem CID: 611); 6-Methyl-5-hepten-2-one (PubChem CID:9862).

Keywords: *Solanum lycopersicum* L., sugar, acid, volatiles, PLS models, consumer panel.

1. Introduction

Consumer complains on tomato flavor became commonplace during the 90s (Bruhn et al., 1991). At the same time, the interest in tomato landraces increased exponentially. Quality markets specialized in high quality produces exploited this interest and their price in the market quadrupled (Cebolla-Cornejo et al., 2007).

Several causes explain the loss of flavor in modern commercial varieties. First, the maximization of the limits on productivity did not contribute to maintaining a high level of sugars in fruit, as photoassimilates have to be partitioned into more sinks, and thus, the accumulation of hexoses per fruit decreases (Bertin et al., 2000). Apart from raising productivity, the need to reduce yield losses led to the introgression of genes, most of them from wild relatives. It has been proven the existence of a considerable level of linkage drag in tomato that can lead to important metabolomic changes (Zhu et al., 2018). During these selection programs, breeders focused on production traits and external appearance and little attention was paid to flavor. Consequently, several favorable alleles regarding the production of aroma volatiles were lost over consecutive breeding cycles (Tieman et al., 2017), and thus, flavor was deeply deteriorated. Additionally, the introgression of new specific traits resulted in negative side effects. It would be the case of delayed fruit ripening, controlled by genes such as *rin*, which decreases consumer acceptance (Causse et al., 2003), probably due to lower carotenoid accumulation and a negative impact on apocarotenoid production, which affect taste perception. The use of genes such as uniform ripening to improve external appearance also had a negative impact on sugar accumulation, as it reduced the accumulation of photoassimilates directly synthesized in fruit (Powell et al., 2012).

In order to establish breeding objectives, tomato flavor started to be modelled from the seventies. Stevens et al. (1977) determined that tomato flavor was deeply influenced by the accumulation of sugars, acids and the relation between them. In this sense, Malundo et al. (1995) established that for a certain sugar content there is an optimal acid accumulation level and increased acid accumulation had a negative effect on preference. Years later, Baldwin et al. (1998) reported that the perception of sweetness is better explained by sucrose equivalent levels which weighs each sugar content by its sweetening power. The contribution to sourness of each organic acid is more complicated. Malic acid may be perceived as sourer (De Bruyn et al., 1971), but the accumulation levels of citric acid are higher. It seems that glutamic acid does not play an important role in sourness, but some studies reported that a high ratio between sucrose equivalents and glutamic acid may be convenient to improve tomato taste (Bucheli et al., 1999; Fulton et al., 2002). Nonetheless, Casals et al. (2018) concluded the opposite, reporting that consumers value glutamic acid contents positively, at least in cherry tomatoes.

The volatile profile of tomato is rather complex and it has an impact on flavor in two ways. The first is the direct impact on aroma. Among the 400 volatiles described in tomato fruit, the initial studies in tomato confirmed that at least 30 had an impact on tomato aroma (Buttery & Ling, 1993). Unlike other species such as banana, in tomato there is not a volatile with a major impact on sensory perception. Additionally, several compounds influence how sweetness and sourness are perceived (Baldwin et al., 1998). They are classified attending to their chemical characteristics (alcohols, aldehydes, apocarotenoids…) or their metabolic origin. Branched chain aminoacid derived volatiles (3- and 2-methylbutanal…) and fatty acid derived volatiles (hexanal, E-2-hexenal, Z-3-hexenal…) tend to offer green notes to tomato aroma and are usually found at high concentrations. Apocarotenoids, which are found at low concentrations, are usually perceived as floral or fruity. On the other hand, phenylalanine derived volatiles, including compounds such as phenylacetaldehyde, 2-phenylethanol, guaiacol or eugenol, have a sensory effect difficult to predict as they have been reported to have both positive and negative effect, probably depending on their different concentrations (Martina et al., 2021). Tandon et al. (2000) suggested that tomato flavor would improve by increasing the levels of compounds contributing to floral (6-methyl-5-hepten-2-one and β-ionone), fruity (cis-3-hexenal and geranylacetone) and fresh (3-methylbutanol and 1-penten-3-one) notes, or by decreasing the levels in compounds which contribute to stale (hexanal, trans-2-hexenal and 3-methylbutanal), pungent (2-isobutylthiazole) and alcohol (2-phenylethanol) notes.

Several classic studies have tried to model flavor perception and chemical composition using linear regression models (i.e. Baldwin et al., 1998; Tandon et al., 2003; Abegaz et al., 2004). In a first approach, some of them have been performed either using a broad spectrum of variance represented in few materials. In contrast, recent approaches have analyzed the variation in large collections of different materials including landraces, commercial cultivars, or even wild representatives. For example, Tieman et al. (2017) did not offer a prediction model, but highlighted the importance of 33 chemicals that correlated with consumer liking and 37 that significantly correlated with flavor intensity, with 28 of them being associated with both overall liking and flavor intensity.

Few studies have been focused on the analysis of specific landraces and few have predicted models using similar materials with subtle differences. But the truth is that consumers highly appreciated tomato landraces and are willing to pay a price premium for them (Cebolla-Cornejo et al., 2007). This is the case of the Spanish Moruno variety of tomato which is highly appreciated by its organoleptic characteristics (Moreno et al., 2019). Tomato landraces are exposed to several factors such as seed-mixing, cross-pollination, and farmer selection (Cortes-Olmos et al., 2015) that increase their variability even in the accumulation of taste-related compounds (Cebolla-Cornejo et al., 2013). Therefore, the analysis of a collection of Moruno populations offers an incredible opportunity to increase our knowledge regarding consumer preference in tomato as conditioned by the accumulation of taste and aroma related compounds.

2. Materials and methods

2.1. Plant materials and field trials

A collection of 30 populations (accessions) belonging to the Spanish "Moruno" tomato landrace (Table 1) highly appreciated by consumers due to its high quality, was analyzed in a 4 year-field study. These populations have an indeterminate growth habit and are characterized by medium to large sized fruits, a dark red or brown colour, strong to medium-ribbing intensity, dark shoulders and a predominantly flattened shape. All of them were obtained from local farmers or from the Spanish seedbanks. The study was developed on 10 different populations per year during the Years 1 to 3; in the Year 4, seven of the populations studied in the previous seasons were considered. These populations were selected considering their best results in the previous trials regarding specific parameters including sensory evaluation. Additionally, the commercial control "Royesta" F1 hybrid (Reimer Seeds), with high acceptance by consumers in Mediterranean areas, was included as a reference.

Cultivation was performed during the spring-summer growing cycle (May to Sept) at the experimental farm of the Research Centre "El Chaparrillo", Regional Institute for Agro-Food and Forestry Research and Development (39º0'N, 3º56'W, altitude 640 m), in Ciudad Real (Central Spain). The climate of this region is continental Mediterranean, with a mean, maximum and minimum air temperatures during the four cropping periods at a range of 20.4 to 22.3°C, 28.3 to 31.0°C and 11.4 to 13.0°C, respectively.

The field trials were conducted in a randomized complete block design with four replicates. Each experimental plot consisted of eight plants (32 plants per population) staked with a separation of 2.0 m between rows and 1.0 m between plants. For the different controls, the central six plants of each plot were considered. Plants were cultivated using organic farming practices (EC n.834/2007), and no chemical fertilizers nor pesticides were applied. Common fertilization and trickle irrigation practices for tomato organic farming production cultivation in the area were followed.

Table 1. Populations evaluated of the "Moruno" tomato landrace and year of cultivation.

Year	Accession	Local name	Origin		
			Town	Province	Coordinates
1	SL-2	"Plano de El Avellanar"	San Pablo de los Montes	Toledo	39°32′N 4°19′W
1	SL-6	"Moruno de San Pablo"	San Pablo de los Montes	Toledo	39°32′N 4°19′W
1	SL-11	"Moruno de El Avellanar"	San Pablo de los Montes	Toledo	39°32′N 4°19′W
1, 4	SL-25	"Moruno"	La Malaguilla	Guadalajara	40°49′N 3°15′W
1	SL-27	"Morado"	Anchuras	Ciudad Real	39°28′N 4°50′W
1, 4	SL-33	"Negrillo"	Almoguera	Guadalajara	40°18'N 2°59′W
1	SL-41	"Negro rosa"	Elche de la Sierra	Albacete	38°27′N 2°3′W
1, 4	SL-62	"Moruno"	Socuéllamos	Ciudad Real	39°17′N 2°47′W
1, 4	SL-72	"Bonito"	Ciudad Real	Ciudad Real	38°59′N 3°55′W
1	SL-74	"Moruno"	Ciudad Real	Ciudad Real	39º0'N, 3º56'W
2, 4	SL-112	"Moruno de Aguas Nuevas"	Aguas Nuevas	Albacete	38°55′N 1°55′W
2	SL-113	"Moruno"	Aguas Nuevas	Albacete	38°55′N 1°55′W
2	SL-114	"Moruno"	Aguas Nuevas	Albacete	38°55′N 1°55′W
2	SL-116	"Moruno"	Aguas Nuevas	Albacete	38°55′N 1°55′W
2	SL-122	"Morao"	Aguas Nuevas	Albacete	38°55′N 1°55′W
2	SL-136	"Morao"	La Poblachuela	Ciudad Real	38°59′N 3°55′W
2	SL-154	"Moruno"	Elche de la Sierra	Albacete	38°27′N 2°3′W
2	SL-160	"Moruno"	Albacete	Albacete	38°59′N 1°51′W
2	SL-163	"Morao"	Arroba de los Montes	Ciudad Real	39°09′N 4°32′W
2	SL-165	"Morado"	Navas de Estena	Ciudad Real	39°29′N 4°31′W
3	SL-20	"Gordo"	Priego	Cuenca	40°27′N 2°19′W
3	SL-140	"Morao"	Arenales de San Gregorio	Ciudad Real	39°18′N 3°01′W
3	SL-143	"Moruno"	Socuéllamos	Ciudad Real	39°17′N 2°47′W
3	SL-149	"Negro"	Riópar	Albacete	38°30′N 2°25′W
3	SL-150	"Negro"	Riópar	Albacete	38°30′N 2°25′W
3, 4	SL-204	"Morao dulce"	Priego	Cuenca	40°27′N 2°19′W
3	SL-207	"Negro plano"	Brihuega	Guadalajara	40°45′N 2°52′W
3	SL-208	"Morao"	Priego	Cuenca	40°27′N 2°19′W
3	SL-209	"Moruno"	Elche de la Sierra	Albacete	38°27′N 2°3′W
3, 4	SL-252	"Moruno"	El Alcornocal	Ciudad Real	40°44′N 3°52′W

2.2. Sampling and basic determinations

The total yield was recorded throughout the whole harvest period (kg plant^{-1}). For further characterization and sensory analysis, fruits in the optimum ripe stage, when the fruit surface was homogenously red colored, were hand-picked in the middle of the harvesting period (first half of September). Three healthy fruits representing the predominant external appearance of each population were taken from each plant (18 fruits per experimental plot, 72 fruits per population) for the different studies. Mean fruit weight (g), number of locules, and fruit dry matter (obtained in an oven set at 70ºC until constant weight and expressed as grams per 100 g fresh weight) were considered. Basic quality parameters were determined including total soluble solids content (SSC), pH, and total titratable acidity. SSC were measured using a digital refractometer ATAGO PR-32 (Atago Co. LTD, Tokyo, Japan) with automatic temperature compensation, which provides values as °Brix. pH was determined using a pH meter, and titratable acidity was quantified by titrating 5 g of tomato paste with 0.1 mol L^{-1} NaOH to pH 8.1 with an automatic sample titrator (TitroMatic 1S- 2B, Crison, Barcelona, Spain). Acidity was expressed as grams of citric acid equivalent per 100 g fresh weight (% citric acid). Each sample was analyzed three times.

2.3. Quantification of sugars, acids, and volatiles

The levels of compounds related to taste and aroma perception were determined in six additional fruits per experimental plot (one fruit per plant, 24 fruits per population). These analyses included the quantification of reducing sugars (fructose and glucose), acids (citric, malic, glutamic), and volatiles related to aroma. For that purpose, fruits were washed with distilled water, homogenized and stored at -80°C until analysis.

Reducing sugars and acids were determined by capillary zone electrophoresis using a P/ACE System MDQ (Beckman Instruments, Fullerton, CA, USA), following the method described by Roselló et al. (2002). Fused silica capillaries (Polymicro technologies, Phoenix, AZ, USA) were used, with a 50 µm internal diameter, 363 µm external diameter, 67 cm total length, and 60 cm effective length. Capillaries were initially conditioned with NaOH, then separation buffer (20 mM 2,6-piridin dicarboxilic acid and 0.1% w:v hexadimethrine bromide, pH = 12.1) was run for 20 minutes at 20°C. Samples were thawed and centrifuged (510 g) for 5 minutes. The upper phase was then diluted (1:10) in deionized water and filtered using 0.2 μm membranes. Samples were injected hydrodynamically for 20 seconds at 0.5 psi. Then separation took place at -25 kV fixed voltage and 20°C. Capillary was rinsed with SDS (60 mM) for 3 minutes at 20 psi between samples, followed by the separation buffer at 20 psi for 3 minutes. Reagents and standards were purchased from Sigma-Aldrich. Each sample was analysed twice and quantification was performed with calibration curves of 5 points and $R^2>0.98$. Results were expressed as g kg^{-1} fresh weight. This method enables the quantification of glutamic acid, as it can be seen in its optimization reported by Cebolla-Cornejo et al. (2012).

The extraction and analysis of volatiles related to aroma were performed as described by Beltran et al. (2006). A tomato sample of 30 g with 5% (w/w) $CaCl_2$ and with the addition of 50 μL of 15 μg mL^{-1} methyl salicylate-D_4 (surrogate/internal standard) were extracted by dynamic head space (purge-and-trap) using SPE Tenax cartridges (Supelco, Sigma-Aldrich Química S.A., Madrid, Spain) and solvent elution. Chromatographic determination was carried out using a Varian CP-3800 gas chromatograph coupled with a mass spectrometry detector (Saturn 4000, Varian). Separation of the analytes was carried out on a 30 m x 0.25 mm DB-5MS (0.25 μm film thickness) Varian capillary column, using helium at 1 mL min^{-1} as carrier gas. The temperature program was as follows: 45°C for 5 min, then raised to 96ºC at a rate of 3°C min^{-1}, then raised to 150°C at a rate of 6°C/min, and finally raised to 240°C at a rate of 30°C/min, with a final isothermal stage of 1.5 min (total chromatographic analysis time of 36 min). Injection of 1 μL (splitless mode, temperature 200°C) was carried out using a Varian 8400 autosampler. The gas-chromatograph was directly interfaced with the ion trap Varian 4000 mass-spectrometer in the external ionization configuration with an electron ionization energy of 70 eV in the positive ion mode. Transfer line temperature was established at 250°C and ion source and trap temperatures were adjusted to 200°C. Quantitation of analytes in the sample extracts was performed using calibration curves using relative areas to internal standard.

Reference aroma compounds were obtained from Sigma-Aldrich Química S.A. (Madrid, Spain; including Supelco and Fluka products) as pure compounds. Stock solutions of the aroma standards at 500 µg L^{-1} were prepared in acetone and stored at -18°C. Working solutions were prepared by volume dilution in diethyl ether-hexane (1:1). The internal standard methyl salicylate-D_4 was of 99.5% purity and was purchased from SigmaAldrich Sigma-Aldrich Química S.A. (Madrid, Spain). Calcium chloride 97% (Riedel de Haen) was purchased from Supelco (Sigma-Aldrich Química S.A., Madrid, Spain). Organic solvents (hexane, ethyl acetate, diethyl ether) of trace residue analysis quality were purchased from Scharlab (Barcelona, Spain). Results were expressed as ng g^{-1} fresh weight.

2.4. Sensory analysis

The remaining fruits were used for the sensory analysis by a consumer panel. This panel consisted of 15 to 20 panellists, depending on the year, including males and females in similar rates, aged 22-52. The panellists were all familiar with taste panel procedures and the terminology used, and consumed tomatoes. The sensory analysis was performed within the day after harvest, and until then the fruits were stored at ~15ºC. Fruits were washed, cut radially into wedges (about eight wedges per fruit), and coded with random numbers. The panellists ranked the fruit samples from the different populations according to the following sensory attributes: sweetness, sourness, taste intensity, aroma intensity, aftertaste persistence, aftertaste agreeability, flesh firmness, skin firmness, grainy texture, floury texture, juicy texture, and overall flavor acceptability. For rating the different fruit textures (grainy, floury, or juicy texture), the percentage of panelists who appreciated each texture is expressed. For the other sensory attributes, a hedonic scale from 1 to 9 (1 = low satisfaction or intensity; 9 = high satisfaction or intensity) was used (Baldwin et al., 2015). The panellists were given water and unsalted crackers between samples, and the tests were conducted in partitioned booths in a climate controlled tasting room.

2.5. Statistical analysis

A Partial Least Square (PLS) regression model for overall flavor acceptability (Y vector) from the other sensory descriptors (X variables) was developed (Supplementary Table 1). PLS regression models with the whole set of samples were also obtained for each sensory descriptor variable (Y vector) from compositional variables (X matrix). SSC, Titratable acidity, pH, malic acid, citric acid, glutamic acid, fructose, glucose, sucrose, sucrose equivalents, Citric acid equivalents and 39 volatile compounds were used as direct compositional variables. Several ratios between sugars and/or acid compounds and inverse or quadratic forms of compositional variables were also included in the in the PLS models (Supplementary Table 2) using a total of 198 X variables in the initial models. Sucrose equivalents were calculated as the weighted sum of sugar concentration using the relative sweetening power of each sugar: 1 for sucrose, 1.73 for fructose and 0.74 for glucose (Baldwin et al., 1998). Citric acid equivalents were calculated as the weight sum of citric and malic acid considering their relative sourness: 1 for citric acid and 1.14 for malic acid (Stevens et al., 1977).

Prior to modelling, data were pretreated using autoscale (mean center and scale each variable to unit standard deviation) to correct different variable scaling and units. Venetian blinds was chosen as cross-validation method to allow an estimation of the model performance. Outlier identification was performed using a graphical evaluation of Q residuals and leverage. Any outlier point that showed a large Q residual or unusual distribution was removed and the model was recalculated. Normalized residuals and leverage parameters were also considered for outlier identification (values < -3 or > 3) and elimination in response variables.

Variable selections were performed to improve the initial PLS models using a multistage criterion. First, an interval PLS (iPLS) forward variable selection procedure (Nørgaard et al., 2000) starting from the variables reported as determinants in tomato flavour (Baldwin et al., 1998; Tandon et al., 2003; Abegaz et al., 2004; Tieman et al. 2017) were executed to find the first set of explaining variables. Second an iPLS reverse variable selection (Nørgaard et al., 2000) from the previous set of selected variables was executed to refine the initial selection, discarding irrelevant variables. A final refinement in the variable selection of the model was performed discarding variables using the selectivity ratio (Sratio) criterion (0.1 fraction removed per iteration, Rajalahti et al., 2009) to obtain the final models (Supplementary Tables 1 and 2).

Resulting prediction model performance was evaluated in terms of outlier diagnostics, the number of latent variables (LV), coefficient of determination of cross-validation (R^2_{CV}), cross-validation bias (CV bias) and root mean square error of cross-validation (RMSECV). The ratios RMSECV/Range of variation were also calculated as a percentage for each descriptor in order to contextualize the results.

All PLS models were performed using Matlab v 9.8 (Mathworks Inc, Natick, MA, USA) and the PLS Toolbox 9.0 for Matlab (Eigenvector Research Inc, Wenatchee, WA, USA). Pairwise correlations were graphically represented as heatmaps using the software heatmapper (http://www.heatmapper.ca). Principal component analysis (PCA) biplots were obtained to describe the variation in volatiles and sensory variables. This analysis was performed with S-plus v.8.01 (Insightful Corp., Seattle, WA, United States).

3. Results and discussion

3.1. "Moruno" landrace: Levels of variation in composition and sensory perception

Spain, as the introduction point of tomato from America into Europe, has a high level of diversity of this species. Centuries of cultivation resulted in a high number of landraces which are highly appreciated by consumers. Accordingly, they are willing to pay up to 4.7 times the price of conventional modern varieties in order to recover the true flavor of tomato (Cebolla-Cornejo et al., 2007). Among these landraces, "Moruno" outstands by the high appreciation of its savory fruits in the area (Moreno et al., 2019). Consequently, it was considered a good case study to model consumer perception of tomato landraces.

"Moruno" tomatoes can be recognized by a common external appearance with big ribbed fruits and dark color, but important differences between populations (accessions) can be found. Indeed, in the present study, the different "Moruno" populations showed a high level of variation in morpho-agronomical traits (Fig. 1a). In general, fruits showed medium to big size, with a high number of locules and dark color, while plant yield remained generally under 10 kg plant^{-1}, but the coefficient of variation (%CV) for fruit weight reached 28%, being even higher for plant yield (44%). Previous studies with a lower number of populations of this landrace had already showed important differences in fruit size and plant yield, as some populations doubled the values reached by others (Moreno et al., 2019). This seems to be a common trend at least in Spanish tomato landraces. In this sense, Cebolla-Cornejo et al. (2013) found high levels of variation not only between populations but within populations of different Spanish landraces.

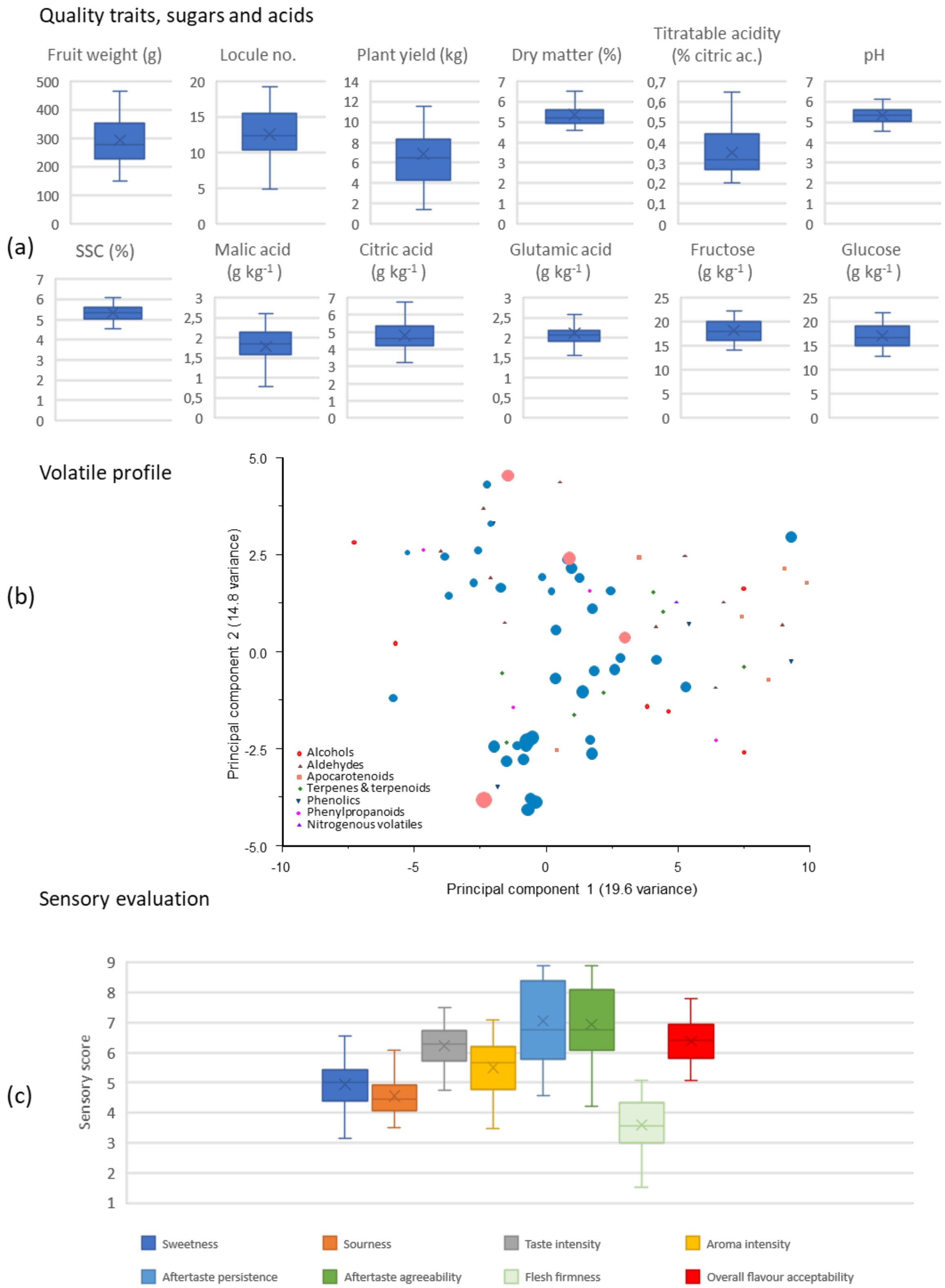

Fig. 1. Profile of variation in the collection of samples assayed regarding to **(a)** agronomical traits, fruit quality parameters and compounds related to taste; **(b)** principal component analysis biplot of fruit volatiles representing the variability the volatile profile related to aroma (red dots: commercial control; blue dots: "Moruno" populations; dot size is proportional to yield); **(c)** sensory scores (1: low satisfaction or intensity; 9: high satisfaction or intensity) of the descriptors evaluated by the consumer panel. Error bars represent standard deviations.

Differences in basic quality parameters including dry matter, titratable acidity, pH, or SSC were more limited (Fig. 1a). pH was the most stable trait with %CV of 2%, followed by SSC (8%) and dry matter (11%), while the variation in titratable acidity was considerably high (30%). When individual sugars and acids were analyzed, a higher level of variation was detected. It was higher in acids than in sugars and especially high in malic acids content (%CV=26%), as expected considering the high variation detected in titratable acidity.

The variation in the volatile profile was schematized using a principal component analysis (PCA) (Fig. 1b), which confirmed a high dispersion of the "Moruno" populations in all the evaluated volatiles classes: alcohols, aldehydes, apocarotenoids, terpenes and terpenoids, phenolics, phenylpropanoids, and nitrogenous volatiles. This variability was also found in the F1 hybrid commercial control, highlighting the considerable effect of the environment on the volatile profile. In this case, it should be considered that the expected variability of commercial F1 hybrids in genotype is negligible, thus, the differences have to be due to environmental effects. Interestingly most of the populations with lower and higher yields plotted in a profile characterized by lower accumulation of most volatiles, specially apocarotenoids. Indeed, these populations present negative scores for principal component 1 and this component presents positive loadings with most volatiles, and especially carotenoids.

The high level of variation in individual compounds associated with taste and aroma perception (Fig. 1a and 1b) was reflected in the values obtained in sensory descriptors evaluated by the consumer panel (Fig. 1c). This variation was higher in traits related to aroma perception and sweetness and lower for sourness (Fig. 1c). A PCA confirmed the sensory variability present in the populations evaluated by the consumer panel (Fig. 2). Environmental effects can be, again, clearly visualized through the dispersion of the F1 hybrid commercial control. Compared to the dispersion observed in the volatile profile (Fig. 1b), the effect of the environment on sensory descriptors is highly reduced (Fig. 2). Consequently, the higher variability observed for sensory perceptions in populations of "Moruno" landrace must be related to genetic differences. In fact, it has been suggested that tomato landraces are affected by seed mixing and spontaneous cross-pollination events (Cortés-Olmos et al., 2015). Then, farmers would apply strong selections to recover fruit morphology traits typical of the landrace. But variation would be maintained in internal traits related to functional or organoleptic quality.

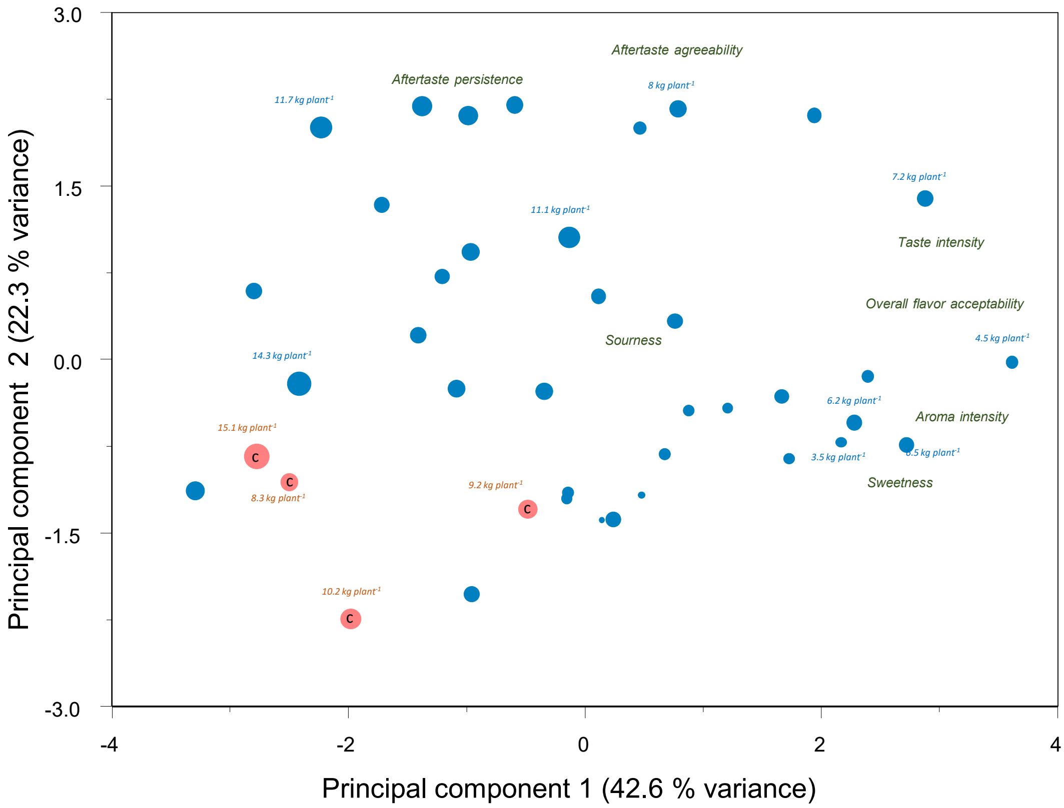

Fig. 2. Biplot representation (scores and loadings) of principal component analysis of sensory descriptors of "Moruno" landrace populations (blue dots), and commercial control (C, red dots). Dot size is proportional to plant yield.

Interestingly, this biplot (Fig. 2) confirmed the difficulty of combining overall flavor acceptability and yield (dot size is proportional to plant yield). In this sense, those accessions with higher yields (close to the best performance of the commercial control) are far from the ideal consumer perception summarized by the overall flavor acceptability. Thus, accessions with a lower yield, generally performed better in sensory terms. Nonetheless, it was still possible to identify accessions with acceptable yields and good sensory performance. Altogether, the high diversity in soluble solids, volatiles, and sensory perception obtained, enabled the development of reliable models of consumer flavor perception.

3.2. Overall flavor perception as affected by other sensory descriptors

In order to analyze which descriptors had a higher influence on consumer acceptability of this high quality landrace, a multi-stage selection variable procedure (iPLS and Sratio methods) was performed. The analysis suggested two latent variables that included sweetness, sourness, taste, aroma, aftertaste persistence, aftertaste agreeability, and flesh firmness as main determinants of overall flavor acceptability (Supplementary Table 1). Other textural perceptions were excluded as they were not significant. The model offered a moderate cross-validation coefficient of determination (R^2_{cv} =0.61), but with a restricted error, with a root mean square error of cross-validation (RMSECV) value of 0.41, representing a 15% of the range of variation of the overall flavor acceptability. PLS overall flavor perception model was exported to a linear model. According to this model, overall flavor acceptability can be inferred from the regression coefficients as all descriptors used the same scale. It depended positively on sweetness, taste and aroma intensities, and negatively of sourness, with similar contributions in absolute values. On the other hand, aftertaste agreeability, aftertaste persistence, and flesh firmness only introduced slight tinges. In recent models using neuronal networks, Cortina et al. (2018) described, in Andean landraces and other materials, a higher preference for tomatoes rated high in sweetness and intermediate in sourness, a preference already reported by Baldwin et al. (1998) in a collection of 24 cultivars, though in that work they later found correlations of overall acceptability with sweetness, but not with sourness. This perception seems to agree with our results.

3.3. Modelling sensory descriptors with compositional variables

Specific compositional PLS models were developed for each sensory descriptor (Supplementary Table 2, Fig. 3). After identifying and removing outliers, iPLS forward models were initially ran considering each sensory descriptor and the concentrations of taste and aroma related compounds, as well as inverse and quadratic derivatives, other derived variables such as ratios between sugars and acids, sucrose equivalents or citric acid equivalents, which had been previously linked to tomato flavor as they weight with the sweetening or acidulant power of each compound (Galiana-Balaguer et al., 2018). In some cases, such as in sweetness, the initially obtained model excluded compounds or variables that had been previously described as important in similar regressions or that showed high linear correlation. Accordingly, the models were repeated forcing the initial inclusion of these variables in the iPLS forward analysis.

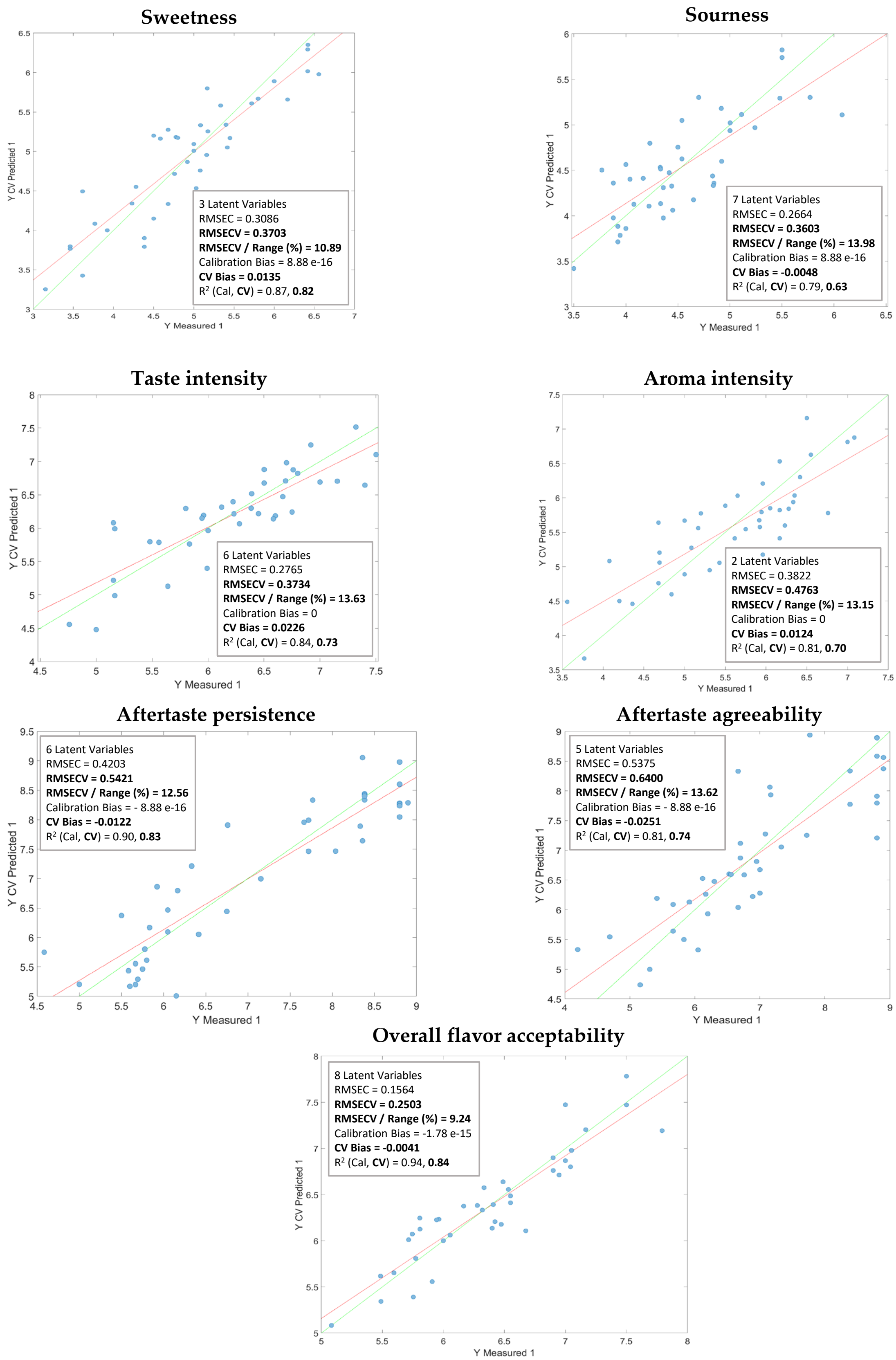

Fig. 3. Performance of PLS models relating sensory descriptors and chemical composition. RMSE: Root Mean Square Error; C: calibration; CV: cross-validation; RMSECV/Range (%): ratio RMSECV to range of variation of the descriptor. Red line: Linear regression model between predicted and measured values. Green line: 1:1 relationship between predicted and measured values.

Although the model performance improved in some cases, it was necessary to check if all the forced variables were really required. For that purpose, a subsequent iPLS reverse analysis was performed to discard spurious variables and the performance increased after deleting some of the initially required variables. Alternative variable selection methods such as VIP (variable importance in projection) or Sratio (selectivity ratio), that focus on variables with higher linear correlation, were also evaluated, but the resulting PLS models showed a worse performance, with higher RMSECV values and lower R^2_{cv}. Accordingly, the models for the different sensory descriptors did not always include variables with high linear correlation. For example, sweetness presented relevant (>|0.4|) positive correlations with SSC, titratable acidity, hexanal, 6-methyl-5-hepten-2-one, and beta-cyclocitral and negative with fructose to glucose ratio, 2-phenylethanol, alpha-pinene, and 3-carene (Supplementary Fig. 1). But the final model did not include most of these compounds. Even when they were initially forced to participate in the model, its performance worsened considerably. In a final stage, the Sratio variable selection method was applied to the selected set of variables to identify variables that could be excluded considering their low contribution to the phenotype. Some of the variables with low mean contribution to the final predicted value were discarded. The final PLS models obtained applying the described multi-stage variable selection procedure for each descriptor were exported to linear models and the mean contribution of each variable to the predicted descriptor value was calculated (Fig. 4). In some cases, variables with low contribution to the descriptor values were included in the model, but their removal decreased both RMSECV and R^2_{cv} values, so they were maintained.

Sweetness perception was finally modelled using three latent variables that included 19 initial variables. The model offered a R^2_{cv} value of 0.82 and a RMSECV of 0.37 was obtained, which represented a 10.89% of the range of variation of the descriptor, %RMSECV (Supplementary Table 2, Fig. 3). SSC, sucrose equivalents and glutamic^{-2} were included in the model, with a large mean contribution to the descriptor of the former (Fig. 4).

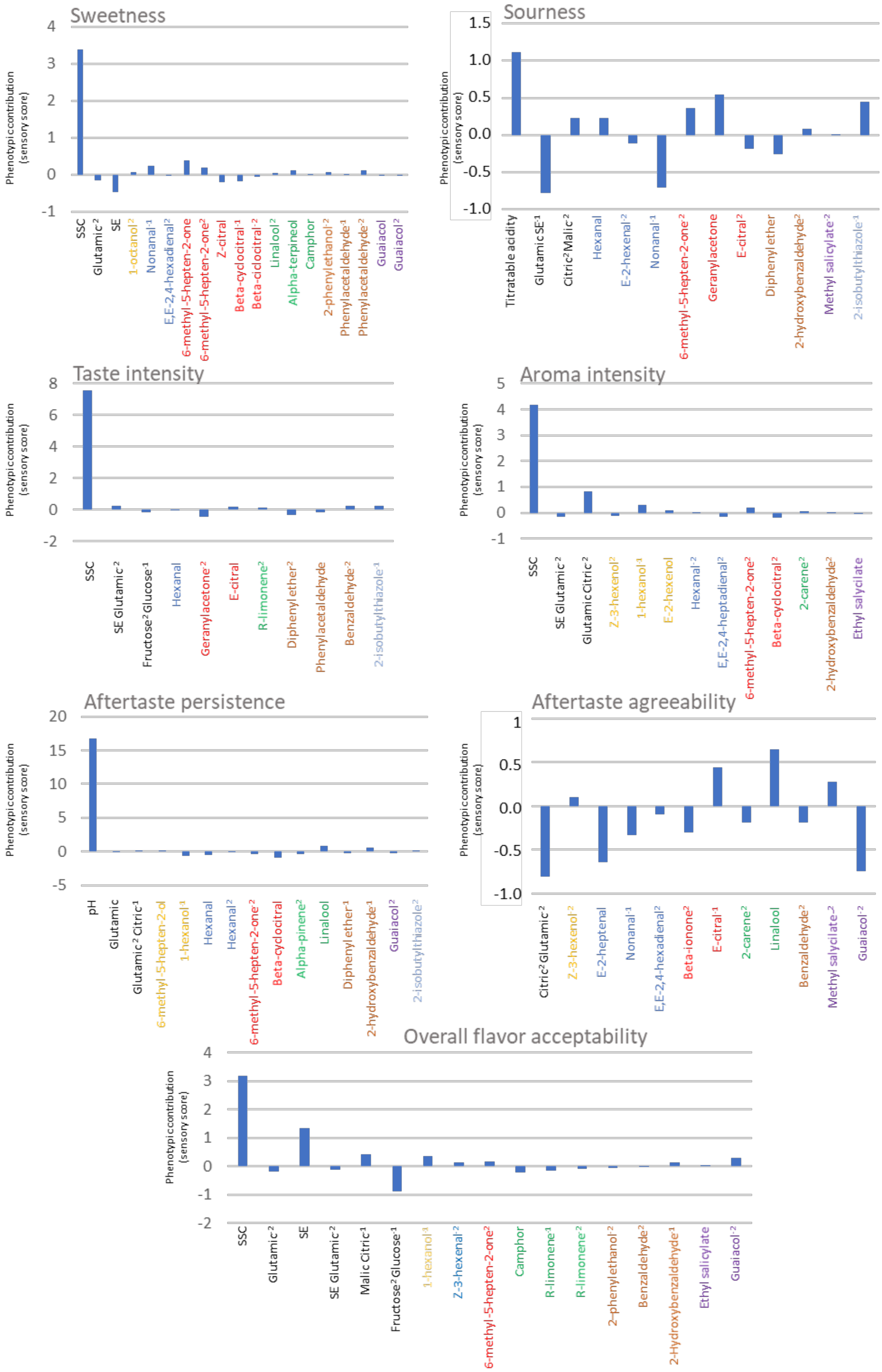

Fig. 4. Mean phenotypic contribution of each compound to the mean predicted value of each descriptor calculated with PLS models. Groups of compounds: sugars and acids, alcohols, aldehydes, phenylpropanoids, apocarotenoids, terpenes & terpenoids, phenolics, nitrogenous volatiles.

The model also included 16 variables related to volatile contents, with minor contributions to the descriptor. Among them, stands out the positive contribution of 6-methyl-5-hepten-2-one. The inverse negative relationship of beta-cyclocitral indicated a positive role, while negative roles were identified in nonanal and z-citral (neral).

The perception of sweetness depends on the amount of sugars present in the fruit. These contents are represented in the model by SSC and sucrose equivalents (SE). The last variable represents the sum of individual sugars weighed by their sweetening power. SE already showed a high correlation with sweetness perception in the models developed by Baldwin et al (1998). Tandon et al. (2003) also found a higher relationship between SE and sweetness than with the use of total sugars. In the model obtained in the present work, SE nuances the contribution of SSC. The participation of acids in the model was expected, but no variable related with acidity was selected in the model. Traditionally, titratable acidity has been linked with the perception of not only sourness but also sweetness (Kader et al., 1977; Tandon et al., 2003). In our case, glutamic^{-2} made slight contributions to the descriptor value, indicating a preference for higher glutamic values. Few studies have analyzed in depth the role of glutamic acid in tomato flavor perception. Bucheli et al. (1999) revealed a major role of the SE glutamic^{-1} ratio, thus suggesting a negative role. But in a more recent study, Casals et al. (2018) confirmed the importance not only of acidity in the perception of sweetness, but also glutamic acid levels in cherry and standard fresh market tomatoes, as in our model.

Several works have remarked the influence of volatiles in the perception of sweetness. In 1998, Baldwin et al. already described the importance of apocarotenoids, offering fruity or floral notes, which were related to this sensory perception, as well as some alcohols, while Krumbein and Auerswald in the same year highlighted the role of 1-penten-3one and 2-methyl-4-pentenal. Later, Tandon et al. (2003) found high correlation levels between sweetness and isobutylthiazole and acetone, and more recently, Baldwin et al. (2015) confirmed the relevance of volatiles such as acetaldehyde, hexanal, trans-2-hexenal, 1-penten-2-one, 6-methyl-5-hepten-2-one, geranylacetone, b-ionone, 2 + 3-methylbutanol, cis-3-hexenol and methylsalicylate in the perception of sweetness. Tieman et al. (2012), in a large studio including commercial varieties and landraces, reported that sweetness was related to fructose, geranial, 2-methylbutanal, 3-methyl-1-butanol and other compounds associated with flavor intensity, including 2-butylacetate, cis-3-hexen-1-ol, citric acid, 3-methyl-1-butanol, 2-methylbutanal, 1-octen-3-one, and E,E-2,4-decadienal. Following a similar approach years later, Tieman et al. (2017) highlighted the role of apocarotenoids in the perception of sweetness, though the analysis of correlations was focused on overall liking and flavor intensity. In addition, not all the studies have included as many volatiles in the perception of sweetness. For example, in a recent study, Cheng et al. (2020) only stood out the role of E,E-2,4-decadienal. In our model, a high number of variables (16) related to 12 volatiles were selected, including one alcohol, two aldehydes, three apocarotenoids, three terpenoids, one phenylpropanoid and two phenolics (Fig. 4). However, the contribution of several of them to the descriptor value was very low, and only phenylacetaldehyde, 6-methyl-5-hepten-2-one, nonanal, z-citral, beta-cyclocitral and alpha-terpineol outstood.

Although apocarotenoids appear consistently related to the perception of sweetness, it seems clear that other volatiles also play an important role. It also can be concluded that considering the disparity in the volatiles selected in different works, specific models for flavor perception would be required for specific materials, considering the divergence between general models generated with a high number of different genotypes and specific models. It seems clear though that 6-methyl-5-hepten-2-one plays a major role. Although each volatile makes a low mean contribution to the descriptor value, altogether represent (in absolute values) a 36.5% of the mean predicted value of sweetness (Fig. 5).

The model for sourness perception included titratable acidity and the ratios glutamic SE^{-1} and $citric^{2}$ $malic^{-2}$ among soluble compounds, as well as 10 variables related to three aldehydes, three apocarotenoids, one phenylpropanoid, two phenolics and one nitrogenous volatile (Fig. 4). Seven latent variables were selected in the PLS model, which offered a moderate performance, with R^2_{cv} = 0.63 and RMSECV of 0.36, representing a 14% of the range of variation of the descriptor (Supplementary Table 2, Fig. 3).

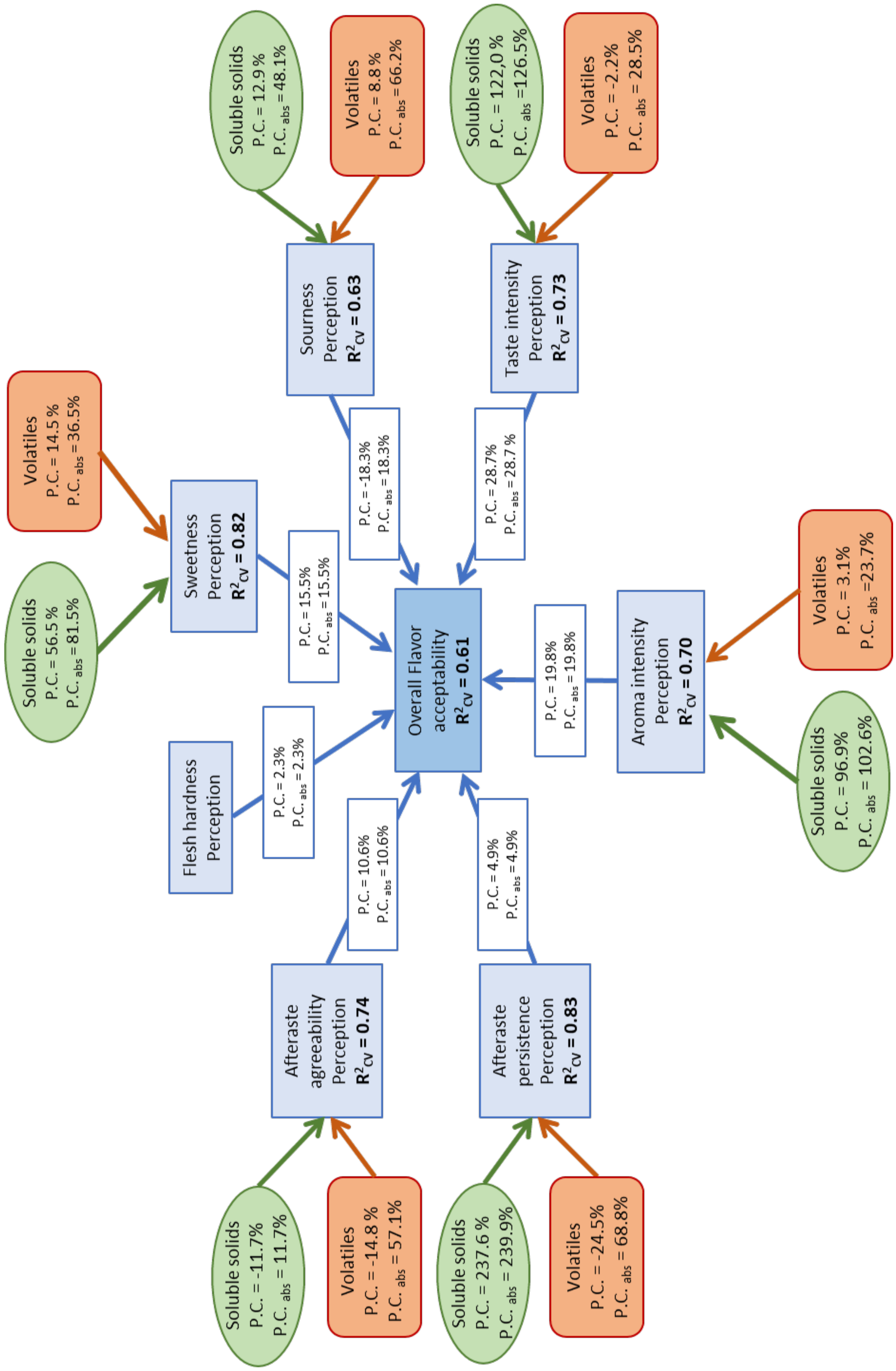

Fig. 5. PLS model relating overall flavor acceptability with other sensory constructs and PLS models relating sensory descriptors and variables related to soluble solids or volatiles contents. P.C.: phenotypic contribution of each group of variables (sum) to the mean predicted value of the descriptor. P.C.$_{abs}$: P.C. calculated with contributions in absolute values. R^2_{cv}: coefficient of determination of the model for cross-valitation.

Volatiles played a major role in the phenotypic contribution of the sourness perception, even higher than soluble solids in absolute values (Fig. 5). Several works have also related not only sugars and acids but also volatiles in the perception of sourness. Baldwin et al. (1998) highlighted the correlations with SSC, pH, acetaldehyde, acetone, 2-isobutylthiazole, geranlyacetone, beta-ionone, ethanol, hexanal and cis-3-hexenal. Tandon et al. (2003) obtained a more limited model including titratable acidity and pH and considered acetone and beta-ionone as positively correlated with sourness. In both models, beta-ionone was selected as an important compound conditioning sourness. However, in our model this compound was not included, but other apocarotenoids did, including geranylacetone with a positive role, and E-citral (geranial) and 6-methyl-5-hepten-2-one with a negative role (the latter due to an inverse relationship). Hexanal was included in the model as in the work by Baldwin et al. (1998). Traditionally, it became clear that not only acids affect the perception of sourness, as their relationship with acids is also crucial, being the tomatoes perceived as more acidic with lower values of hexoses. Interestingly, volatiles also affect sourness perception. In some models, the same compound has been related both with sweetness and sourness. This effect has also been found with nonanal^{-1} in both models. In the case of 6-methyl-5-hepten-2-one, though, higher values contributed positively to sweetness but negatively to sourness.

It is difficult to compare the rest of the descriptors with previous works, as they are not usually included. In our case, the model for taste intensity showed a good performance (Supplementary Table 2, Fig. 3). Six latent variables were selected, offering R^2_{cv} = 0.73 and RMSECV = 0.37 (13.6% of the range of variation). Sugars seemed to play an important role, represented with the selection of the variables SSC, SE glutamic^{-2}, and fructose^2 glucose^{-1} ratios (Fig. 4), with a major contribution of the former. But also one aldehyde, two apocarotenoids, one terpenoid, three phenolic and one nitrogenous volatiles were included (Fig. 4). Nonetheless, in this case, the phenotypic value of the descriptor was mainly determined by soluble solids (Fig. 5).

In the aroma intensity model (Supplementary Table 2, Fig. 3), two latent variables were selected, offering a moderate performance, with R^2_{cv} = 0.70 and RMSECV = 0.48 (13.2% of the range of variation). As expected, a higher number of volatiles (10) was included in the model, with three alcohols, two aldehydes, two apocarotenoids, one terpenoid, one phenylpropanoid and one phenolic (Fig. 4). But interestingly, soluble solids had a major contribution *via* SSC, SE glutamic^{-2} and glutamic citric^{-2}. In fact, the mean phenotypic contribution in real or absolute values of soluble solids was higher than that of the volatiles (Fig. 5).

Aftertaste descriptors offered higher RMSECV values (Supplementary Table 2, Fig. 3), though they represented less than 13.6% of the range of variation. Specifically, aftertaste persistence was modelled with six latent variables (RMSECV = 0.54 and R^2_{cv} = 0.83) and included the variables pH, glutamic, the ratio glutamic^2 citric$^{-1,}$ and 12 variables related to 11 volatiles: two alcohols, one aldehyde two apocarotenoids, two terpenoids, and three phenylpropanoids (Fig. 4). Interestingly, pH made with difference the highest mean contribution to the descriptor value, with fruits with higher pH offering a higher aftertaste persistence. On the other hand, the model for aftertaste agreeability (Supplementary Table 2, Fig. 3) required five latent variables (R^2_{cv} = 0.74, RMSECV = 0.64, %RMSECV = 13.6%). The model selected the ratio citric^2 glutamic^{-2} and 11 volatiles: one alcohol, three aldehydes, two apocarotenoids, two terpenoids, two phenylpropanoids and one phenolic volatile (Fig. 4). Interestingly, aftertaste persistence was mainly conditioned by contributions of soluble solids variables while aftertaste agreeability depended more on volatile variables (Fig. 5).

In the study by Baldwin et al. (1998), aftertaste intensity was found to correlate with acetaldehyde, beta-ionone and ethanol, though more compounds were related to aftertaste sour (acetaldehyde, hexanal, and 2-isobutylthiazole) and aftertaste bitter (soluble solids, acetaldehyde, l-penten-3-one, beta-ionone, ethanol, methanol, and 2+ 3-methylbutanol). Among them, only 2-isobutylthiazole was selected in our case for aftertaste persistence with a positive role and beta-ionone for aftertaste agreeability but with a negative role. It seems necessary then to obtain specific models for specific materials.

Finally, overall flavor acceptability was modelled with eight latent variables and offered an excellent fit with R^2_{cv}=0.84 and RMSECV=0.25, representing less than 10% of the range of variation (Supplementary Table 2, Fig. 3). Soluble solids represented the main mean contributions to the descriptor value, with high positive contributions of SSC and SE and negative of the ratio fructose^{2} glucose^{-1}. Glutamic^{-2} and the ratios SE glutamic^{-2} and malic citric^{-1} were also included in the model, but their contributions to the phenotypic value were lower (Fig. 4).

Baldwin et al. (1998) concluded that among soluble solids components, overall acceptability was related to the ratio total sugars to titratable acidity, SE to titratable acidity and titratable acidity. Previously Malundo et al. (1995) found that increasing sugars and acids improved tomato acceptability but only up to a point, at which higher acid levels reduced liking, justifying the importance of sugar to acids ratio. In our case, with a specific type of tomato, the role of sugar is clear with the contributions of SSC and SE to acceptability, while the role of acids is represented by the malic citric^{-1} ratio. Interestingly, our model highlights the role of glutamic acid and the ratio SE glutamic^{-2}, suggesting a beneficial effect of increasing concentrations of glutamic acid considering the negative relationship of these variables with acceptability. Bucheli et al. (1999) found that the best markers for tomato fruitiness in tomato varieties included reducing sugars, malic and glutamic acid, with a negative role of the latter. Similarly, in *S. pimpinellifolium* breeding lines they observed a positive role of reducing sugars to glutamic acid ratio, reducing sugars, glucose, and a negative one of glutamic acid. This was one of the first mentions of a relative role of glutamic acid, which the authors justified considering that the most effective activity as flavor potentiator of this compound was exerted at a pH (5.5-8.0) higher than tomato pH (4.0-4.6). Our results revisit the role of glutamic acid, highlighting the fact that, in certain contexts, glutamic acid can play a positive role in flavor acceptability. Eleven variables related to 10 volatiles were also selected in our model. In absolute values, the contributions of volatiles to the overall flavor acceptability values (1.58) represented a 26% of the contribution of variables related to soluble solids (6.1). Considering inverse relationships, according to the model, acceptability mainly increased with increasing contents of 6-methyl-5-hepten-2-one and R-limonene and decreasing levels of 1-hexanol, camphor, 2-hydroxybenzaldehyde and guaiacol, some of them due to an inverse relationship. According to *The good scent company* database (www.thegoodscentscompany.com), positive volatiles represent fruity and citric notes, while negative volatiles represent pungent green (1-hexanol), medicinal (camphor and 2-hydroxybenzaldehyde) and spicy woody phenolic notes (guaiacol). Additionally, Z-3-hexanal^{-2}, 2-phenylethanol^{-2}, benzaldehyde^{2} and ethyl salycilate were also included in the model but with lower contributions.

Baldwin et al. (1998) found that E-6,10-dimetyl-5,9-undecadien-2-one and 6-methyl-5-hepten-2-one had preferable odors in tomato fruits. Years later, Piambino et al. (2012) found in a collection of different tomato types that Z-3-hexen-1-ol and 2-isobutylthiazole played a major positive role, while 2-butanol, benzylalcohol, 6-methyl-5-hepten-2-ol and Z-2-nonenal had a negative influence. In a latter review, Klee and Tiemann (2018) highlighted the positive role on consumer preferences of fructose and glucose, 1-nitro-2-phenylethane, 1-nitro-3-methylbutane, 1-penten-3-one, 2,5-dimethyl-4-hydroxy-3(2H)-furanone, 2-isobutylthiazole, 2-phenylethanol, 3-pentanone, 6-methyl-5-hepten-2-ol, 6-methyl-5-hepten-2-one, benzaldehyde, benzyl cyanide, Z-4-decenal, heptaldehyde, isovaleric acid, isovaleronitrile, nonyl aldehyde, phenylacetaldehyde, E-2-heptenal, E-2-pentenal, E-3-hexen-1-ol, and a negative role of butyl acetate, eugenol, hexyl acetate, isobutyl acetate, prenyl acetate and salicylaldehyde. One of the last studies on this topic developed by Cheng et al. (2020) in a collection of different tomato types, reported that the volatiles that contributed more to tomato flavor were E,E-2,4-decadienal, E-2-hexenal, 1-(2,6,6-trimethyl1-cyclohexen-1-yl)-2-buten-1-one, 6-methyl-5-hepten-2-one, hexanal, 2-isobutylthiazole, 2,6,6-timethyl-1-cyclohexene1-carboxaldehyde, E-6,10-dimetyl-5,9-undecadien-2-one, 4-allyl-2-methoxyphenol, E-2-heptenal, E-2-octenal, Z-3-hexen-1-ol, and methyl salicylate. They also found that malic acid, 2-E-3-(3-pentyl-2-oxiranyl)acrylaldehyde, 2-hydroxy-ethyl benzoate, methyl salicylate and 2-methoxyphenol were disliked by the evaluation panels. All being considered, it seems evident that floral and fruity notes of apocarotenoids are generally preferred, while pungent and medicinal notes might be negative. But the role of each specific compound changes depending on the different studies. This seems to point out again that specific models are required for specific materials and that models generated with a wide diversity of tomato types may not be as generalizable as it would have been considered. This problem has also been detected in other contexts related to tomato quality. For example, NIR models developed with a high number of different varieties representing a wide spectrum of variation are not as reliable as specific models developed for specific contexts (Ibáñez et al., 2019)

4. Conclusions

Models relating sensory evaluation and fruit composition in tomato, despite offering a general view, are rather variable, probably due to the differences in the materials tested. Even global models, developed with a high number of varieties grown in different years, seem to do not apply to specific contexts. Indeed, the soluble solids and volatiles identified as determinants for flavor perception in the present study showed only a minimum overlap with global scale studios. Our results confirm the role of volatiles in the definition of descriptors related to taste, as well as the contribution of soluble solids to aroma perception. In terms of overall acceptability, tough, soluble solids play a major role that is then tinged by different volatiles. Among them our work shows that 6-methyl-5-hepten-2-one is a key volatile that should be specially considered in the development of breeding programs. In this context, the specific accumulation of individual sugars, should also be addressed, as they tinge sensory descriptors. Nonetheless, it seems clear that general models do not apply in the case of tomato and it would be required to develop specific models for specific materials. In this context, this study offers valuable information from the methodological point of view. PLS methods have proved to be a reliable tool to model tomato sensory perception, offering low %REMSECV values and highlighting the role of certain compounds that do not outstand for high linear correlations. The models obtained already represent ideal targets to be considered in the development of breeding programs, considering the high acceptability of the tomato landrace “Moruno”. On the other hand, the evaluation of the sensory profile of different populations of the same landrace again highlights the huge variation present in these genetic resources. It has become commonplace the generalization that landraces outstand by their organoleptic characteristics, but the truth is that numerous populations of landraces have lost during their evolution their valued flavor. It is necessary to continue with depuration programs to consolidate their presence in high quality markets and their on-farm conservation. But in the process, it should be considered that extremely high yields would highly impact the sensory profile of tomatoes.

Credit authorship contribution statement:

Jaime Villena: methodology & formal analysis, review & editing. **Carmen Moreno**: conceptualization, methodology, review & editing, supervision & project administration. **Marta M. Moreno**: conceptualization, methodology, review & editing, supervision & project administration. **Joaquín Beltrán**: methodology & formal analysis. **Salvador Roselló**: methodology & formal analysis, visualization, review and editing. **Jaime Cebolla-Cornejo**: visualization, writing original draft, writing review & editing.

Declaration of competing interests: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments: This research has been funded by the National Institute for Agricultural and Food Research and Technology (INIA), Ministry of Science and Innovation (grant code: RTA 00083-C02-01). The authors thank *'Servei Central d'Instrumentació Científica (SCIC) de la Universitat Jaume I'* for the use of scientific analytical equipment.

Data availability: 10.5281/zenodo.6963114

Supplementary materials: Supplementary material associated with this article can be found, in the online version.

References:

Abegaz, E. G., Tandon, K. S., Scott, J. W., Baldwin, E. A., Shewfelt, R. L. 2004. Partitioning taste from aromatic flavor notes of fresh tomato (*Lycopersicon esculentum*, Mill) to develop predictive models as a function of volatile and nonvolatile components. Postharvest biology and technology, 34, 227-235. https://doi.org/10.1016/j.postharvbio.2004.05.023.

Baldwin E. A., Scott, J. W., Einstein, M., Malundo, T. M. M., Carr, B. T., Shewfelt, R. L., Tandon, K. S. 1998. Relationship between sensory and instrumental analysis for tomato flavor. Journal of the American society for Horticultural Science, 123, 906-915. https://doi.org/10.21273/JASHS.123.5.906.

Baldwin, E. A., Scott, J. W., Bai, J. 2015. Sensory and chemical flavor analyses of tomato genotypes grown in Florida during three different growing seasons in multiple years. Journal of the American Society for Horticultural Science, 140, 490-503. https://doi.org/10.21273/JASHS.140.5.490.

Beltran, J., Serrano, E., López, F. J., Peruga, A., Valcarcel, M., Rosello, S. 2006. Comparison of two quantitative GC–MS methods for analysis of tomato aroma based on purge-and-trap and on solid-phase microextraction. Analytical and bioanalytical chemistry, 385, 1255-1264. https://doi.org/10.1007/s00216-006-0410-9.

Bertin, N., Guichard, S., Leonardi, C., Longuenesse, J. J., Langlois, D., Navez, B. 2000. Seasonal evolution of the quality of fresh glasshouse tomatoes under Mediterranean conditions, as affected by air vapour pressure deficit and plant fruit load. Annals of Botany, 85, 741-750.

Bruhn, C. M., Feldman, N., Garlitz, C., Harwood, J., Ivans, E., Marshall, M., Riley, A., Thurber, D., Williamson, E. 1991. Consumer perceptions of quality: apricots, cantaloupes, peaches, pears, strawberries, and tomatoes. Journal of food quality, 14, 187-195. https://doi.org/10.1111/j.1745-4557.1991.tb00060.x.

Bucheli, P., Voirol, E., Torre, R. R., Lopez, J., Rytz, A., Tanksley, S. D., Petiard, V., de la Torre, R. 1999. Definition of nonvolatile markers for flavor of tomato (*Lycopersicon esculentum* Mill.) as tools in selection and breeding. Journal of Agricultural and Food Chemistry, 47, 659-664. https://doi.org/10.1021/jf980875l.

Buttery, R. G., Ling, L. C. 1993. Volatile components of tomato fruit and plant parts: relationship and biogenesis. In R. Teranishi, R. G. Buttery, H. Sugisawa (Eds.), Bioactive volatile compounds from plants (pp. 22–34). ACS Symposium Series Nr. 525. Washington, D.C.: American Chemical Society.

Casals, J., Rivera, A., Sabaté, J., Romero del Castillo, R., Simó, J. 2018. Cherry and fresh market tomatoes: differences in chemical, morphological, and sensory traits and their implications for consumer acceptance. Agronomy, 9, 9. https://doi.org/10.3390/agronomy9010009.

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, 2342-2350.

Cebolla-Cornejo, J., Soler, S., Nuez, F. 2007. Genetic erosion of traditional varieties of vegetable crops in Europe: tomato cultivation in Valencia (Spain) as a case study. International Journal of Plant Production, 1, 113-128. https://doi.org/10.22069/IJPP.2012.531.

Cebolla-Cornejo, J., Valcárcel, M., Herrero-Martínez, J. M., Roselló, S., Nuez, F. 2012. High efficiency joint CZE determination of sugars and acids in vegetables and fruits. Electrophoresis, 33, 2416-2423.

Cebolla-Cornejo, J., Roselló, S., Nuez, F. 2013. Phenotypic and genetic diversity of Spanish tomato landraces. Scientia Horticulturae, 162, 150-164.
Cortés-Olmos, C., Valcárcel, J. V., Roselló, J., Díez, M.J., Cebolla-Cornejo, J. 2015. Traditional Eastern Spanish varieties of tomato. Scientia Agricola, 72, 420-431. http://dx.doi.org/10.1590/0103-9016-2014-0322.
Cortina, P. R., Santiago, A. N., Sance, M. M., Peralta, I. E., Carrari, F., Asis, R. 2018. Neuronal network analyses reveal novel associations between volatile organic compounds and sensory properties of tomato fruits. Metabolomics, 14, 1-15. http://dx.doi.org/10.1007/s11306-018-1355-7.
Cheng, G., Chang, P., Shen, Y., Wu, L., El-Sappah, A.H., Zhang, F., Liang, Y. 2020. Comparing the flavor characteristics of 71 tomato (*Solanum lycopersicum*) accessions in central Shaanxi. Frontiers in Plant Science, 11, 1882.
De Bruyn, J. W., Garretsen, F., Kooistra, E. 1971. Variation in taste and chemical composition of the tomato (*Lycopersicon esculentum* Mill.). Euphytica, 20, 214-227.
Fulton, T. M., Bucheli, P., Voiol, E., Lopez, J., Petiard, V., Tanksley, S. D. 2002. Quantitative trait loci (QTL) affecting sugars, organic acids and other biochemical properties possibly contributing to flavor, identified in four advanced backcross populations of tomato. Euphytica, 127, 163-167. http://dx.doi.org/10.1023/a:1020209930031.
Galiana-Balaguer, L., Ibanez, G., Cebolla-Cornejo, J., Rosello, S. 2018. Evaluation of germplasm in *Solanum* section *Lycopersicon* for tomato taste improvement. Turkish Journal of Agriculture and Forestry, 42, 309-321. http://dx.doi.org/10.3906/tar-1712-61.
Ibáñez, G., Cebolla-Cornejo, J., Martí, R., Roselló, S., Valcárcel, M. 2019. Non-destructive determination of taste-related compounds in tomato using NIR spectra. Journal of Food Engineering, 263, 237-242. https://doi.org/10.1016/j.jfoodeng.2019.07.004.
Kader, A. A., Stevens, M. A., Morris, L., Algazi, M. 1977. Effect of fruit ripeness when picked on flavor and composition in fresh market tomatoes [Quality]. Journal American Society for Horticultural Science, 112, 724-731.
Klee, H. J., Tieman, D. M. 2018. The genetics of fruit flavour preferences. Nature Reviews Genetics, 19, 347-356. https://doi.org/10.1038/s41576-018-0002-5.
Malundo, T. M. M., Shewfelt, R. L., Scott, J. W. 1995. Flavor quality of fresh tomato (*Lycopersicon esculentum* Mill.) as affected by sugar and acid levels. Postharvest Biology and Technology, 6, 103-110. https://doi.org/10.1016/0925-5214(94)00052-T.
Martina, M., Tikunov, Y., Portis, E., Bovy, A. G. 2021. The genetic basis of tomato aroma. *Genes*, 12, 226.
Moreno, M. M., Villena, J., González-Mora, S., Moreno, C. 2019. Response of healthy local tomato (*Solanum lycopersicum* L.) populations to grafting in organic farming. Scientific Reports, 9, 1-10. https://doi.org/10.1038/s41598-019-41018-2.

Nørgaard, L., Saudland, A., Wagner, J., Nielsen, J. P., Munck, L., Engelsen, S. B. 2000. Interval partial least-squares regression (i PLS): A comparative chemometric study with an example from near-infrared spectroscopy. Applied spectroscopy, 54, 413-419. https://doi.org/10.1366/0003702001949500.

Piombino, P.; Sinesio, F.; Moneta, E.; Cammareri, M.; Genovese, A.; Lisanti, M.T.; Mogno, M.R.; Paparaio, M.; Termolino, P.; Moio, L.; Grandillo, S. 2012. Investating physicochemical, volatile, and sensory parameters playing a positive or a negative role on tomato linking. Food Res. Int., 50, 409-419. https://doi.org/10.1016/j.foodres.2012.10.033.

Powell, A. L., Nguyen, C. V., Hill, T., Cheng, K. L., Figueroa-Balderas, R., Aktas, H., Ashrafi, H., Pons, C., Fernández-Muñoz, R., Vicente, A., Lopez-Baltazar, J. 2012. Uniform ripening encodes a Golden 2-like transcription factor regulating tomato fruit chloroplast development. Science, 336, 1711-1715. https://doi.org/10.1126/science.1222218.

Rajalahti, T., Arneberg, R., Berven, F. S., Myhr, K. M., Ulvik, R. J., Kvalheim, O. M. 2009. Biomarker discovery in mass spectral profiles by means of selectivity ratio plot. Chemometrics and Intelligent Laboratory Systems, 95, 35-48. https://doi.org/10.1016/j.chemolab.2008.08.004.

Roselló, S., Galiana-Balaguer, L., Herrero-Martínez, J. M., Maquieira, A., Nuez, F. 2002. Simultaneous quantification of the main organic acids and carbohydrates involved in tomato flavour using capillary zone electrophoresis. Journal of the Science of Food and Agriculture, 82, 1101-1106. https://doi.org/10.1002/jsfa.1153.

Stevens, M. A., Kader, A. A., Algazi, M. 1977. Genotypic variation for flavor and composition in fresh market tomatoes. Journal American Society for Horticultural Science, 102, 680-689.

Tandon, K. S., Baldwin, E. A., Shewfelt, R. L. 2000. Aroma perception of individual volatile compounds in fresh tomatoes (*Lycopersicon esculentum* Mill.) as affected by the medium of evaluation. Postharvest Biology and Technology, 20, 261-268. https://doi.org/10.1016/S0925-5214(00)00143-5.

Tandon, K. S., Baldwin, E. A., Scott, J. W., Shewfelt, R.L. 2003. Linking sensory descriptors to volatile and nonvolatile components of fresh tomato flavor. Journal of Food science, 68, 2366-2371. https://doi.org/10.1111/j.1365-2621.2003.tb05774.x.

Tieman, D., Bliss, P., McIntyre, L. M., Blandon-Ubeda, A., Bies, D., Odabasi, A. Z., Rodríguez, G. R., van der Knaap, E., Taylor, M. G., Goulet, C., Mageroy, M. H. 2012. The chemical interactions underlying tomato flavor preferences. Current Biology, 22, 1035-1039. https://doi.org/10.1016/j.cub.2012.04.016.

Tieman, D., Zhu, G., Resende, M. F., Lin, T., Nguyen, C., Bies, D., Rambla, J. L., Beltran, K. S. O., Taylor, M., Zhang, B., Ikeda, H. 2017. A chemical genetic roadmap to improved tomato flavor. Science, 355, 391-394. https://doi.org/10.1126/science.aal1556.

Zhu, G., Wang, S., Huang, Z., Zhang, S., Liao, Q., Zhang, C., Lin, T., Qin, M., Peng, M., Yang, C., Cao, X. 2018. Rewiring of the fruit metabolome in tomato breeding. Cell, 172, 249-261. https://doi.org/10.1016/j.cell.2017.12.019.