Artificial neural networks and genetic dissimilarity among saladette type dwarf tomato plant populations

Camila Soares de Oliveira a, Gabriel Mascarenhas Maciel b,*, Ana Carolina Silva Siquieroli c, Danilo Araújo Gomes a, Nadia Mendes Diniz a, José Magno Queiroz Luz b, Rickey Yoshio Yada d

a Institute of Agricultural Sciences, Federal University of Uberlândia, Uberlândia, Brazil
b Institute of Agricultural Sciences, Federal University of Uberlândia, Monte Carmelo, Brazil
c Institute of Biotechnology, Federal University of Uberlândia, Monte Carmelo, Brazil
d Faculty of Land and Food Systems, University of British Columbia, Vancouver, BC, Canada

A R T I C L E   I N F O

Keywords:
Solanum lycopersicum
Dwarfism
Backcrossing
Computational intelligence
Genetic dissimilarity

A B S T R A C T

Studies have shown that dwarf plants have the potential for use in obtaining hybrids. The aim of this study was to evaluate the agronomic potential and genetic dissimilarity of saladette type dwarf tomato plant populations through the use of artificial neural networks (ANNs). The following traits were analyzed: mean fruit weight, transverse and longitudinal fruit diameter, fruit shape, pulp thickness, locule number, internode length, soluble solids content, and β-carotene, lycopene, and leaf zingiberene contents. A dendrogram obtained by the unweighted pair-group method with arithmetic mean (UPGMA) and Kohonen self-organizing maps (SOM) agreed in the distinction of the BC2F3 populations from the dwarf donor parent. SOM was more consistent in identifying the genetic similarities among the BC2F3 dwarf tomato plant populations and allowed for the determination of weights of each variable in the cluster formation. The UFU SD1 13-1 BC2F3 population was revealed to be a promising option for obtaining saladette type dwarf tomato plant lines.

1. Introduction

Tomato is one of the main vegetable crops produced and consumed worldwide (Maham et al., 2020). The market demand for fresh tomatoes with enhanced taste and cooking versatility has expanded the cultivation of saladette/roma tomato plants (Shirahige et al., 2010). However, high production costs, along with the strong susceptibility of the crop to various types of biotic and abiotic stresses have driven the search for alternative, more feasible growth varieties for producers (Wamser et al., 2012; Almeida et al., 2015).

Plant breeding has been an important strategy for increasing crop yield in a profitable and sustainable manner (Barbosa et al., 2011). In cherry/grape tomatoes, internode reduction through breeding with a dwarf parental line (Maciel et al., 2015) has led to promising results in obtaining hybrids with compact plant architectures and high yields (Finzi et al., 2017). Dwarf tomato plant populations have previously been cultured for round/salad-type tomatoes intended for in natura consumption (Finzi et al., 2020), but similar dwarf plants have not yet been identified for the saladette/roma tomato segment.

Characterizing and analyzing the genetic dissimilarity of populations are essential for distinguishing divergent and promising genotypes (Maciel et al., 2018). Traditionally, multivariate techniques have been used, such as dendrograms, the Tocher grouping method, canonical variables, and principal components to characterize the germplasms of normal tomato plants (Maciel et al., 2018; Alsamir et al., 2019; Peixoto et al., 2020) and dwarf tomato varieties (Finzi et al., 2020).

Greater difficulties arise in characterizing and analyzing the germplasm of dwarf tomato plants (Finzi et al., 2017; Finzi et al., 2020), as they have considerably different morphologies than normal-sized tomatoes, which requires the refinement of methods and the use of new optimization techniques. Artificial neural networks (ANNs) have been used to obtain self-organizing maps (SOMs), which may be beneficial alternatives for the study of genetic dissimilarity in saladette type dwarf tomatoes.

These ANNs have comparative advantages in relation to the traditional methodologies, such as the enabling of a non-parametric approach that tolerates loss of data and effectively recognizes patterns and establishes clusters (Kavzoglu and Mather, 2003). SOMs are a class
of ANN that use a competitive learning mechanism through distance as an activation function to recognize similarities between input patterns and establish groupings (Cruz and Nascimento, 2018). This allows for the study of genetic dissimilarity in different germplasms such as rice (Santos et al., 2019), alfalfa (Santos et al., 2020), and cotton (Cardoso et al., 2021).

Thus, the aim of this study was to assess the agronomic potential and genetic dissimilarity of saladette/roma type dwarf tomato plant populations using computational intelligence (ANNs).

2. Material and methods

2.1. Plant material and experimental design

The experiment was conducted from October 2019 to March 2020 at the vegetable crop experimental station of the Universidade Federal de Uberlândia (UFU), Monte Carmelo Campus, MG, Brazil (18° 42’43.19” S, 47°29’55.8” W; altitude 873 m) in an arch-type greenhouse (7 × 21 m) covered with a 150 µm transparent polyethylene film with ultraviolet radiation protection and anti-aphid white screen lateral curtains.

The dwarf tomato plant populations (BCF2) used in this study were obtained from the tomato germplasm bank of UFU. They were derived from two self-fertilizations of a first backcross after hybridization of a pre-commercial homoygous line (UFU MC TOM 5) with a saladette/roma type fruit (recurrent parent) × a dwarf plant line (UFU MC TOM1) with cherry/grape type fruit (Maciel et al., 2015).

Nineteen populations were evaluated from this germplasm bank (1: UFU SDI 11-4, 2: UFU SDI 17-8, 3: UFU SDI 7-4, 4: UFU SDI 17-5, 5: UFU SDI 4-3, 6: UFU SDI 5-4, 7: UFU SDI 17-7, 8: UFU SDI 13-1, 9: UFU SDI 13-2, 10: UFU SDI 17-1, 11: UFU SDI 13-3, 12: UFU SDI 7-2, 13: UFU SDI 11-5, 14: UFU SDI 10-5, 15: UFU SDI 4-6, 16: UFU SDI 17-9, 17: UFU SDI 17-6, 18: UFU SDI 18-1, and 19: UFU SDI 6-1), along with the donor parent (UFU MC TOM1), the recurrent parent (UFU MC TOMS), and the commercial hybrid Pizzadoro, for a total of 22 treatments.

The genotypes were sown in polystyrene trays (200 cells) filled with a coconut fiber-based substrate (Bioplant Ltda, Nova Ponte-MG, Brazil). Seedlings were transplanted 40 days after sowing (DAS) in 5 L plastic pots containing the same substrate. Crops were treated as recommended for growing tomato plants in a protected environment (Alvarenga, 2013).

The experiment was conducted in a randomized block design (RBD) with four replications, and each experimental plot was composed of six plants distributed in double rows at a spacing of 0.3 × 0.3 m. A spacing of 0.8 m was used between the double rows.

2.2. Agronomic evaluation of fruits

At 90 days after sowing, the fruit from each plot was harvested, counted, and weighed to determine the mean fruit weight (MW). Fifteen tomatoes per plot were subsequently taken as samples, and the following traits were evaluated: transverse fruit diameter (TD), measured from the pedicel scar to the blossom end of the tomato fruit; longitudinal fruit diameter (LD), measured in the transverse direction of the cut fruit; fruit shape (FS), determined by the ratio between the TD and LD of the tomato; and pulp thickness (PT), determined by the greatest distance (thickness) of the mesocarp of the fruit. The number of locules (NL) was determined by directly counting the locules in the tomato.

2.3. Internode length evaluation

Internode length (IL) was calculated as the ratio between the height and the number of plant nodes measured at the end of the crop cycle (155 DAS).

2.4. Fruit quality evaluation

The soluble solids content (SS) was measured using a digital pocket refractometer (Atago PAL-1 3810; Atago Co. Ltd., Tokyo, Japan) in °Brix.

Extraction and quantification of β-carotene (BCC) and lycopene (LC) were conducted according to the methodology adapted from Rodriguez-Amaya (2001). The tomato pulp was ground and 1 g of the material obtained was placed in a glass vial containing 3 mL of 100% acetone (Dinâmica Ltda, Indaiatuba-SP, Brazil). The samples were kept in the dark at a temperature of 8 °C for 48 h. The supernatant was then evaluated by spectrophotometry (Tecnal Ltda, Piracicaba-SP, Brazil), and the absorbance values were obtained for BCC and LC at wavelengths of 450 nm and 470 nm, respectively. Pigments were quantified according to the protocols described by Rodriguez-Amaya (2001) and Rodriguez-Amaya and Kimura (2004).

2.5. Extraction and quantification of zingiberene

The leaf zingiberene content (ZGB) was determined at 80 DAS using a sample composed of eight leaf disks (equivalent to 4.2 cm²) from each plant in the plot. The disks were collected from leaflets from the upper third of the plants and placed in test tubes. The allelochemical zingiberene was extracted and quantified following the methodology described by Freitas et al. (2000).

2.6. Statistical analysis

Analysis of variance was performed using the F-test (α = 0.05). The mean values were compared using the Scott-Knott test (α = 0.05) and Dunnett’s test (α = 0.05), with the dwarf plant donor line (UFU MC TOM1) considered the control for the purpose of determining the gains obtained by backcrossing. The following genetic parameters were also analyzed: genotypic coefficient of variation (h²), genetic coefficient of variation (CVg), and the ratio between the genetic and environmental coefficient of variation (CVg/CVe).

2.7. Multivariate analysis and self-organising map by artificial neural networks

The genetic dissimilarity among the populations was studied using the conventional method by obtaining the Mahalanobis generalized distance matrix. Genetic diversity was represented by the dendrogram obtained by the unweighted pair-group method with arithmetic mean (UPGMA) hierarchical method, validated by the cophenetic correlation coefficient (CCC). In addition, an analysis was performed using computational intelligence (ANN).

The SOM was obtained by applying a non-supervised traditional approach according to the characteristics evaluated and the requirements of the study. Network training to obtain the SOM was performed in 5000 epochs (equal to the total number of comparisons made), with four neurons in each dimension and a pattern of radius neighbor = 1.

The model was validated using different configurations for the number of neurons. Combinations were tested by varying the number of rows (2–5) and columns (2–5). To select the best architecture, 5000 training sessions were conducted for each combination. Thus, it was observed that the combination that best represented the genetic dissimilarity of the dwarf tomato germplasm of the saladette type was that of four rows and four columns (16 neurons) with a neighboring radius pattern of 1, hexagonal neighbor topology, feedforward network architecture with an input layer (means) and an output neuron, and a Euclidean distance-type activation function.

For the determination of the SOM classes, there was competition between the output neurons, with the Euclidean distance as a discriminant function for each input vector. The neuron with the highest
discriminant function value was deemed the winning neuron. Thus, each genotype was allocated to its most representative neuron. Subsequently, in which the closer neurons respond, in the manner of similar input patterns.

All analyses were performed using GENES software, integrated with the R and MATLAB software (Cruz, 2016).

3. Results

3.1. Univariate analysis

3.1.1. Agronomic characterization

The genotypes under study differed in agronomic traits, fruit quality, and leaf zingiberene content (F test, α = 0.05) (Table 1).

The BC$_1$F$_3$ dwarf populations produced larger tomatoes than those of the UFU TOM 1 donor parent (Scott-Knott and Dunnett, α = 0.05). The UFU-SDI 13-1, UFU-SDI 17-1, and UFU-SDI 18-1 populations were notably for the production of tomatoes with MWs greater than 30 g. The UFU-13-1 and UFU-17-1 populations were superior to the other populations, with LDs equal to 5.49 cm and TDs of 3.43 and 3.75 cm, respectively.

The recurrent parent had a mean of 3.38 locules per tomato, while the donor parent had a mean of two units per tomato. The tomatoes of the populations showed intermediate NLs compared to the parents, ranging from 2.28 to 2.85 units. However, the UFU SDi 11-4, UFU SDi 17-8, and UFU SDi 17-9 populations did not differ statistically from the donor parent by the Dunnett test (α = 0.05), and revealed the lowest number of locules among the populations.

The PT of tomatoes from the dwarf plant populations was 155% greater than that of the tomatoes from the donor parent. In general, 42.1% of the populations had a PT greater than 0.55 cm.

An elongated FS was predominant among the BC$_1$F$_3$ populations, parents, and the commercial control. Despite statistical differences, all genotypes had an FS index greater than 1.

The normal phenotype plants had internodes greater than 6.5 cm, whereas the ILs of the BC$_1$F$_3$ dwarf plant populations ranged from 1.66 to 2.44 cm. An IL shorter than 2 cm was found in 47.4% of the population. The UFU SDi 7-4, UFU SDi 4-3, and UFU SDi 13-3 populations did not differ statistically from the donor parent with an IL of 1.31 cm (Dunnett α = 0.05).

In the present study, the UFU TOM 1 donor parent had a greater ZGB (0.22 nm) than the other genotypes in both tests (Scott-Knott and Dunnett, α = 0.05). Lower contents of this allelochemical were observed in cv. Pizzadoro (0.03 nm), the recurrent parent, and the UFU SDi 17-5, UFU SDi 4-3, UFU SDi 13-3, UFU SDi 11-5, UFU SDi 17-6, and UFU SDi 6-1 populations. Intermediate contents of this allelochemical were observed in the UFU SDi 7-4, UFU SDi 5-4, UFU SDi 13-1, UFU SDi 7-2, UFU SDi 10-5, and UFU SDi 17-9 populations.

The expression increases in the BC$_1$F$_3$ populations compared to the donor parent in relation to the agronomic traits attest to the effectiveness of the first backcross (Fig. 1).

3.1.2. Fruit quality characterization

The fruit quality parameters for the BC$_1$F$_3$ dwarf populations, parents, and the commercial cultivar differed for all traits evaluated by the Scott-Knott test (α = 0.05) (Table 2).

The greatest SS (7.10 °Brix) was observed in the fruit produced by the UFU TOM 1 donor parent. The UFU SDi 4-3 and UFU SDi 17-9 populations, together with cv. Pizzadoro predominated, expressing SSs greater than 5.70 °Brix. The UFU SDi 7-4, UFU SDi 13-2, UFU SDi 17-1, and UFU SDi 13-3 populations expressed SSs less than 5.0 °Brix.

The fruit from the UFU TOM 1 donor parent had a high SS, but also a considerable BCC (2.39 mg/100 g). A BCC similar to that of the donor parent was observed in 76.2% of the BC$_1$F$_3$ dwarf populations. In addition, 68.4% of the BC$_1$F$_3$ populations produced fruit with LC greater than 2.71 mg/100 g and equaled the UFU TOM 1 donor parent in both tests (Dunnett and Scott-Knott α = 0.05). The UFU SDi 6-1, UFU SDi 11-5, UFU SDi 13-1, UFU SDi 17-8, UFU SDi 17-7, UFU SDi 17-6, and UFU SDi 18-1 populations had notably high content of both carotenoids.

Except for the LC and BCC, all variables exhibited h$^2$ values greater than 0.70, and CV$_g$/CV$_e$ ratios greater than 1 (Tables 1 and 2).

3.2. Multivariate analysis

3.2.1. Hierarchical cluster analysis (UPGMA)

The dendrogram obtained by the UPGMA (Fig. 2) had a CCC of 0.88 and a distortion of 16.67.

The cophenetic correlation coefficient was established by abrupt level changes in the dendrogram at 8% dissimilarity, by which five groups were retrieved. The UFU TOM 1 donor parent, UFU MC TOM 5 recurrent parent, and cv. Pizzadoro formed three distinct groups. The dwarf populations were divided into two groups, one of which was formed by 64.2% of the BC$_1$F$_3$ populations and the other by the genotypes UFU SDi 17-1, UFU SDi 18-1, and UFU SDi 13-1.

3.2.2. Kohonen self-organizing maps (ANNs/SOM)

Using the SOM method, of the 16 neurons established with the four rows and four columns for the command, the 22 genotypes were classified into 12 classes (Fig. 3).

Classes 1 and 3 consisted of three genotypes and classified the largest number of accessions. Classes 2, 7, 10, and 13 held only one genotype each, showing the genetic dissimilarity of these accessions in relation to the remainder. Classes 4, 6, 8, 9, 11, and 12 each contained two genotypes. No genotype was allocated in Class 5, revealing the low similarity between the genotypes clustered in Classes 1 and 10, and those in Classes 2 and 10. The remaining classes (14, 15, and 16) contained no genotypes.

The UFU SDi 13-3, UFU SDi 11-5, and UFU SDi 13-1 populations and the UFU TOM 1 donor parent formed isolated groups of low similarity with the other populations by the SOM method (Table 3).

The groups that gathered the largest number of populations in representation of SOM were constituted by the following BC$_1$F$_3$ populations: UFU SDi 7-2, UFU SDi 10-5, and UFU SDi 4-6 (Group I); and UFU SDi 11-4, UFU SDi 17-5, and UFU SDi 13-2 (Group III). Among the 12 clusters that were formed in the SOM method, only that constituting the donor parent corroborated the representation of the dendrogram created by the UPGMA method.

Greater genetic dissimilarity was found in the UFU SDi 13-1 population than in the others, since this genotype was allocated to an isolated group corresponding to Class 10 in the SOM (Fig. 3), and to the group with the lowest number of dwarf populations in the dendrogram obtained by the UPGMA method (Fig. 2).

The effect of each variable on each group formed in the SOM was shown in the representation of the neuron topology of the network generated by means of the weights and the association of each input variable with the output neuron (Fig. 4).

The determining traits in distinguishing the UFU TOM 1 donor parent were FS, SS, BCC, and ZGB. The LC was an important trait in identifying the clusters corresponding to Classes 11 and 12, which held the UFU SDi 6-1, UFU SDi 7-4, UFU SDi 17-8, and UFU SDi 17-7 populations. MW, TD, and LD, similar to PT, NL, and LC, were relevant for distinguishing the group formed by the UFU SDi 13-1 population.

4. Discussion

It has been reported that tomato plant characteristics should be prioritized to obtain future cultivars, especially with regard to the various biotic (Ferrero et al., 2020; Zanin et al., 2021) and abiotic
Table 1
Agronomic traits evaluated in 19 BC F3 tomato plant populations, the recurrent parent, donor parent, and commercial control. Monte Carmelo, MG, Brazil, 2020.

| Genotypes   | MW   | LD  | TD  | FS   | PT   | NL  | IL  | ZGB |
|-------------|------|-----|-----|------|------|-----|-----|-----|
| UFU Sdi 11-4 | 17.38 | 4.59 | 2.75 | 1.67 | 0.46 | 2.28 | d   | 2.36 | b   | 0.10 | c   |
| UFU Sdi 17-8 | 21.62 | 4.86 | 2.98 | 1.63 | 0.51 | 2.34 | d   | 2.16 | c   | 0.08 | c   |
| UFU Sdi 7-4  | 25.50 | 4.78 | 3.08 | 1.56 | 0.55 | 2.45 | c   | 1.66 | d   | 0.17 | b   |
| UFU Sdi 17-5 | 20.39 | 4.85 | 3.00 | 1.62 | 0.49 | 2.45 | c   | 2.07 | c   | 0.06 | d   |
| UFU Sdi 4-3  | 23.35 | 4.84 | 3.05 | 1.59 | 0.54 | 2.44 | c   | 1.73 | d   | 0.06 | d   |
| UFU Sdi 5-4  | 26.31 | 5.18 | 3.15 | 1.65 | 0.57 | 2.47 | c   | 1.94 | d   | 0.16 | b   |
| UFU Sdi 17-7 | 18.54 | 4.88 | 2.96 | 1.65 | 0.52 | 2.57 | c   | 2.03 | c   | 0.09 | c   |
| UFU Sdi 13-1 | 41.89 | 5.49 | 3.75 | 1.46 | 0.76 | 2.65 | b   | 2.02 | c   | 0.13 | b   |
| UFU Sdi 13-2 | 19.33 | 4.81 | 2.79 | 1.73 | 0.46 | 2.53 | c   | 1.96 | d   | 0.11 | c   |
| UFU Sdi 17-1 | 30.22 | 5.49 | 3.43 | 1.60 | 0.60 | 2.47 | c   | 1.96 | d   | 0.10 | c   |
| UFU Sdi 13-3 | 22.73 | 4.79 | 3.00 | 1.60 | 0.45 | 2.74 | b   | 1.71 | d   | 0.05 | d   |
| UFU Sdi 7-2  | 25.20 | 4.66 | 3.08 | 1.52 | 0.51 | 2.85 | b   | 1.91 | d   | 0.16 | b   |
| UFU Sdi 11-5 | 16.46 | 4.53 | 2.86 | 1.59 | 0.46 | 2.45 | c   | 2.08 | c   | 0.05 | d   |
| UFU Sdi 10-5 | 22.71 | 4.05 | 2.82 | 1.44 | 0.44 | 2.58 | c   | 1.89 | d   | 0.14 | b   |
| UFU Sdi 4-6  | 26.37 | 4.81 | 3.10 | 1.56 | 0.50 | 2.82 | b   | 1.82 | d   | 0.10 | c   |
| UFU Sdi 17-9 | 19.36 | 5.14 | 2.90 | 1.78 | 0.49 | 2.30 | d   | 2.44 | b   | 0.12 | b   |
| UFU Sdi 17-6 | 22.03 | 5.00 | 3.02 | 1.66 | 0.48 | 2.47 | c   | 2.40 | b   | 0.04 | d   |
| UFU Sdi 18-1 | 33.86 | 5.75 | 3.22 | 1.79 | 0.62 | 2.73 | c   | 2.18 | c   | 0.08 | c   |
| UFU Sdi 6-1  | 25.68 | 4.66 | 3.08 | 1.51 | 0.56 | 2.48 | c   | 2.31 | b   | 0.07 | d   |
| UFU MC TOM5  | 74.34 | 7.53 | 4.48 | 1.69 | 0.94 | 3.38 | a   | 6.93 | a   | 0.05 | d   |
| Pizzadoro    | 55.43 | 6.33 | 4.30 | 1.50 | 0.95 | 2.78 | b   | 6.79 | a   | 0.03 | d   |
| UFU TOM 1    | 5.00  | 1.85 | 1.87 | 0.21 | 2.00 | 1.31 | e   | 0.22 | a   |      |

Mean values followed by different letters in the column differ from each other by the Scott-Knott test at 0.05. *Mean values in the column differ from the UFU MC TOM 1 dwarf donor line control by the Dunnett test at a level of 0.05. h: genotypic coefficients of determination; CVg: genetic coefficients of variation; CVg/Cv: ratio between genetic and environmental coefficients of variation.

MW: mean fruit weight (g); LD: longitudinal diameter (cm); TD: transverse diameter (cm); FS: fruit shape; PT: pulp thickness (cm); NL: number of locules (locules per fruit); IL: internode length (cm); ZGB: zingiberene content (270 nm).
stressors (Oliveira et al., 2021; Wen et al., 2021), as well as the nutritional quality of the fruit (Londoño-Giraldo et al., 2020; Gomes et al., 2021; Oliveira et al., 2022) for a healthy diet (Asensio et al., 2019). A major obstacle to this is the narrow genetic base of the tomato germplasm (Hassan et al., 2021).

Several studies have sought to increase genetic variability and gene introgression by using interspecific crosses of wild species (Peixoto et al., 2020; Dariva et al., 2021). However, the biggest challenge has been in recovering the agronomic potential and nutritional quality of the fruits after each crossing with wild species (Peixoto et al., 2020; Dariva et al., 2021).

In order to conduct the introgression of genes of interest for different types of biotic stress and fruit nutritional quality, and promote higher productivity, research with dwarf tomato plants has been intensified (Maciel et al., 2015; Finzi et al., 2017; Finzi et al., 2020; Cavasin et al., 2021; Gomes et al., 2021; Oliveira et al., 2021). The use of dwarf plants has provided multiple advantages, and obtaining saladette/roma dwarf tomato populations is a promising technology for the future development of higher yielding hybrids, similar to that observed with cherry/grape tomatoes (Finzi et al., 2017).

The potential of dwarf plants is clear for the provision of gene introgression aimed at several agronomic, morphological, and nutritional advantages and a broad spectrum of pest resistance (Tables 1 and 2). Plants with shorter internodes have been used in breeding programs for various crops, and improvements have been reported in yield and plant architecture when a dwarf parent is used (Finzi et al., 2017; Wu et al., 2018; Cho et al., 2021). The dwarf tomato plant populations obtained in this study exhibited short internodes and could be viable alternatives for increasing the yield of future hybrid tomato plants, and thereby obtain plants with more compact architecture that will facilitate management and harvest activities (Frasca et al., 2014; Sun et al., 2019).

The superiority in the size of the tomatoes of the BC₁F₃ dwarf populations compared to those produced by the donor parent confirms the success of the breeding method used in this study. The increase in tomato fruit size is associated with increases in MW, TD, LD, PT, and NL (Marques et al., 2019; Tijskens et al., 2020). Similar results were observed by Finzi et al. (2020), who evaluated the same agronomic variables and reported the superiority of the BC₁F₃ dwarf populations of round type tomatoes in relation to the dwarf donor parent belonging to the cherry/grape segment.

Elongated fruit were predominant among the BC₁F₃ populations, parents, and commercial control, since all genotypes had a FS index greater than 1. The increase in fruit size and the FS index (LD/TD > 1.5) corroborate the standard exhibited by the saladette/roma type tomato fruit (Andrade et al., 2014).

Fruit firmness is generally related to a smaller NL and greater PT (Rodrigues et al., 2010; Amaral Júnior et al., 2017). Thus, the UFU SDi 13-1 population proved to be promising for obtaining firmer fruit.

Resistance to arthropod pests by antixenosis and antibiosis mechanisms has been correlated with high ZGB levels in tomato plant leaflets (Rezende et al., 2020; Zanin et al., 2021). The BC₁F₃ populations had ZGBs that were intermediate to that of the parents. According to Oliveira et al. (2020a), this response is indicative of incomplete dominance of gene activity for this trait. The superiority of ZGB in 68.4% of BC₁F₃ populations of dwarf tomato in relation to the commercial cultivar Pizzadoro reveals the potential of this germplasm for the resistance to a broad spectrum of pests. There is great interest in the introgression of genes related to the presence of this allelochemical in cultivated tomatoes (Rezende et al., 2020). The plants of the populations obtained in this study can be considered excellent alternatives to increase productivity (Finzi et al., 2017) and resistance to pests (Zanin et al., 2021), and minimize the period of the crop genetic improvement program.
The saladette/roma tomato is prominent among the traditional fresh tomato segments, as it has an SS that ranges from 4 to 5 °Brix, leading to a better taste and sweeter flavor of the tomato fruit (Ikeda et al., 2013; Schwarz et al., 2013). In the present study, the SS of the fruit produced by the BC$_1$F$_3$ dwarf populations was similar to that of the recurrent parent and the Pizzadoro commercial cultivar that belongs to the saladette/roma segment. In addition, the mean values expressed for this trait corroborate the SS found in saladette/roma tomato hybrids, showing that the populations are promising for this segment (Andrade et al., 2014).

An increase in LC and BCC in tomato fruit is one of the aims of crop breeding programs (Hazra et al., 2018; Londono-Giraldo et al., 2020; 2021). Therefore, the characterization of germplasm regarding the levels of these carotenoids can provide relevant information to achieve this aim. In this study, LC and BCC were expressive and consistent, corroborating the values obtained by Bhandari et al. (2021) and Asensio et al. (2019), respectively.

The change in eating habits toward an adequate caloric intake and efficient nutritional scheme has stimulated research into the development of biofunctional foods for improved health and reduction of disease risk (Stajčić et al., 2015; Asensio et al., 2019; Athinodorou et al., 2021). The populations BC$_1$F$_3$: UFU SDi 11-4, UFU SDi 17-8, UFU SDi 17-7, UFU SDi 17-6, and UFU SDi 18-1 evaluated in this study are promising for aggregating these traits in a

### Table 2

| Genotypes          | SS   | LC     | BCC   |
|--------------------|------|--------|-------|
| UFU SDi 11-4       | 5.01c*| 2.77 a | 1.61b |
| UFU SDi 17-8       | 5.18c*| 3.55 a | 2.24 a|
| UFU SDi 7-4        | 4.50 d*| 3.52 a | 1.72 b|
| UFU SDi 17-5       | 5.26c*| 2.91 a | 1.61b |
| UFU SDi 4-3        | 6.28b*| 2.75 a | 1.75b |
| UFU SDi 5-4        | 5.05 c*| 3.07 a | 1.61 b|
| UFU SDi 17-7       | 5.10 c*| 2.07 a | 1.61 b|
| UFU SDi 13-1       | 5.40 c*| 3.34 a | 1.99 a|
| UFU SDi 13-2       | 4.62 d*| 1.89 b | 1.61 b|
| UFU SDi 17-1       | 6.46 d*| 2.87 a | 1.61 b|
| UFU SDi 13-3       | 4.32 d*| 2.27 b | 1.20 b*|
| UFU SDi 7-2        | 5.25 c*| 1.88 b | 1.49 b*|
| UFU SDi 11-5       | 5.18 c*| 2.88 a | 1.61 b|
| UFU SDi 10-5       | 5.32 c*| 2.79 a | 1.61 b|
| UFU SDi 4-6        | 5.45 c*| 2.41 b | 1.37 b*|
| UFU SDi 17-9       | 5.58 b*| 2.60 b | 1.84 a|
| UFU SDi 17-6       | 5.10 c*| 3.09 a | 1.98 a|
| UFU SDi 18-1       | 5.55 c*| 3.12 a | 2.05 a|
| UFU SDi 6-1        | 4.93 c*| 3.38 a | 1.91 a|
| UFU MC TOM5        | 5.21 c*| 2.43 b | 1.65 b|
| Pizzadoro          | 5.73 b*| 2.63 b | 1.49 b*|
| UFU TOM1           | 7.10 a | 2.71 a | 2.39 a|
| Mean               | 5.28  | 2.77   | 1.78  |
| %CV                | 6.86  | 22.57  | 28.53 |
| h$^2$              | 90.73 | 54.06  | 51.87 |
| CV$\beta$          | 10.74 | 12.54  | 13.15 |
| CV$\beta$/CV$\epsilon$ | 1.56 | 0.54   | 0.51  |

SS: soluble solids content (°Brix); LC: lycopene content (mg/100 g); BCC: β-carotene content (mg/100 g). *Mean values followed by different letters in the column differ from each other by the Scott-Knott test at 0.05. Mean values in the column differ from the UFU MC TOM1 dwarf donor line control by the Dunnett test at a level of 0.05. h$^2$: genotypic coefficients of determination; CV$\beta$: genetic coefficients of variation; CV$\beta$/CV$\epsilon$: ratio between genetic and environmental coefficients of variation.
saladette/roma type dwarf tomato plant breeding program to obtain a functional food rich in antioxidants, given its superiority over the recurring parent and the commercial cultivar Pizzadoro. Saladette/Italian tomato cultivars have gained popularity in the tabletop consumption market because of their high SS, pleasant aroma, PT, and texture (Shirahige et al., 2010; Andrade et al., 2014).

However, there are few studies on LC and BCC for the fruits of this segment. The populations that presented greater contents of these compounds than the Pizzadoro cultivar expressed an average of 2.09 and 3.20 mg/100 g for BCC and LC, respectively. Seabra Junior et al. (2022) evaluated six commercial Italian tomato cultivars and reported BCC and LC below 1.33 mg/100 g and 2.73 mg/100 g, respectively. This demonstrates that saladette-type dwarf tomato plants have the potential to increase crop productivity (Finzi et al., 2017) and improve the fruit quality of future hybrids and cultivars.

One of the advantages of using the Kohonen network in relation to other multivariate techniques is the ability to determine the weights of the variables in the formation of the groups and visualize the related distance ($D^2$) matrix (Cruz et al., 2012). Through this method, genetic dissimilarity was found between the BC$_1$F$_3$ populations, parents, and the commercial cultivar by the formation of five groups. This methodology is traditionally used to represent genetic dissimilarity among tomato plant populations in various segments (Maciel et al., 2018; Finzi et al., 2020; Peixoto et al., 2020).

SOMs are obtained by computational intelligence tools through neural networks, and they are promising technologies in the study of genetic dissimilarity among populations (Santos et al., 2019). Such maps allow the visualization of similar patterns and classification of data based on the distances between them (Oliveira et al., 2020b). The hexagonal topology used in this study allowed for a better arrangement of the neurons and a minimization of possible errors in the classification process (Kohonen, 2014). The organization of the topological structure reflects the similarity among the genotypes under study, allowing for classification by approximation (Santos et al., 2019; Gomes et al., 2021).

The SOM had a greater discriminating power among the genotypes compared to the UPGMA method. While the UPGMA clustering method allocated the BC$_1$F$_3$ dwarf populations into two groups, the SOM distributed them among ten groups. The better capacity of the SOM in discriminating the genotypes in relation to the UPGMA method was also reported by Cardoso et al. (2021), who compared the two methods in the evaluation of genetic dissimilarity of colored cotton genotypes. Gomes et al. (2021) classified salad-type dwarf tomato populations and also observed a greater number of groups by the SOM method than the UPGMA method, and determined that this methodology was more efficient for studying the genetic dissimilarity in dwarf tomato germplasm.

The differences found in the clustering performed by the two methodologies used in this study can be explained by the fact that the dendrogram obtained by $D^2$ takes the means and variances of the traits under study into consideration in the clustering process (Cruz et al., 2012). However, in the Kohonen method, clustering is not affected by experimental errors, and this method has a strong ability to simulate neural networks, which amplify the input data and estimate new values based on different synaptic weights for each neuron, and organize the groups by order of proximity based on Euclidean distance (Oliveira et al., 2020b).

One of the advantages of using the Kohonen network in relation to other multivariate techniques is the ability to determine the weights of the variables in the formation of the groups and visualize the related

Table 3

| Groups | Classes | Genotypes |
|--------|---------|-----------|
| I      | 1       | UFU SDI 7-2; UFU SDI 10-5; UFU SDI 4-6 |
| II     | 2       | UFU SDI 13-3 |
| III    | 3       | UFU SDI 11-4; UFU SDI 17-5; UFU SDI 13-2 |
| IV     | 4       | UFU SDI 17-9; UFU SDI 4-3 |
| V      | 5       | UFU SDI 17-6; UFU SDI 18-1 |
| VI     | 7       | UFU SDI 7-7 |
| VII    | 8       | UFU SDI 17-7 |
| VIII   | 9       | UFU MC TOM 5; cv. Pizzadoro |
| IX     | 10      | UFU SDI 13-1 |
| X      | 11      | UFU SDI 5-4; UFU SDI 17-1 |
| XI     | 12      | UFU SDI 17-6; UFU SDI 18-1 |
| XII    | 13      | UFU SDI 7-2; UFU SDI 10-5; UFU SDI 4-6 |

Fig. 4. Traits and weights in the activation of each SOM neuron. Lighter colors represent greater effect of a variable on the group determined by the neuron.
correlations (Cardoso et al., 2021; Gomes et al., 2021). The similarity in the patterns of colors observed for the MW, LD, and TD, and PT traits emphasizes their correlations, showing that these traits are closely related to fruit size (Tijssen et al., 2020).

The traits that distinguished the donor parent (UFU TOM 1) were FS, SS, BBC, and ZGB. In means comparison testing, this genotype showed superior results for these traits. The association between the univariate emphasis on these traits is closely related to fruit size (Tijskens et al., 2020).

5. Conclusions

The SOM was more consistent in distinguishing genetic similarity among the BC2F3 populations of saladette/roma type dwarf tomato plants, resulting in a larger number of clusters. The UFU SDI 13-1 BC2F3 dwarf population revealed agronomic potential and superior fruit quality, and is a promising population for the saladette/roma type dwarf tomato plant breeding program.

Declaration of Competing Interest

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

The authors thank the Federal University of Uberlândia (UFU), the Minas Gerais State ResearchFoundation (FAPEMIG), the National Council for Scientific and Technological Development (CNPq), and the Coordination for the Improvement of Higher Education Personnel (CAPES) for funding this research.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.fochms.2021.100056.

References

Almeida, V. S., Silva, D. J. H., Gomes, C. N., Antonio, A. C., Moura, A. D., & Lima, A. L. R. (2015). Sistema Viçosa para o cultivo de tomateiro. Horticultura Brasileira, 33, 74–79. https://doi.org/10.1590/S0100-204X2015000300012

Ahnahm, M., Ahmad, N., Arief, V., Mahmoud, T., & T Efthimesian, R. (2019). Phenotypic diversity and marker-trait association studies under heat stress in tomato (Solanum lycopersicum L.). Aust. J. Crop Sci., 13, 578–587. https://doi.org/10.21475/ajcs.19.13.04.p1581

Alvaranga, M. A. R. (2013). Tomate: produção em campo, em casa-de-vegetação e em hidroponia (2ª ed.). Livraria: Editora UFV, 640 p.

Amaral Júnior, A. T., Graça, A. J., Vivas, M., Viana, A. P., & Rodrigues, R. (2017). Agronomic performance of mini-tomato hybrids from dwarf lines. Ciência e Agrotecnologia, 4, 15–21. https://doi.org/10.1590/1984-70332015v21n0p18

Finzi, R. R., Maciel, G. M., Perez, H. G., Silva, M. F., Peixoto, J. V. M., & Gomes, D. A. (2020). Agronomic potential of BC2F3 dwarf round tomato populations. Ciência e Agrotecnologia, 44, Article e028819. https://doi.org/10.15405/actasciagron.v44i0.32629

Hasson, Z., Ullah, S., Khan, A. A., Shahzad, U., Khurshid, M., Bakhsh, A., … Manzoor, Z. (2021). Phenotypic characterization of exotic tomato germplasm: An excellent breeding resource. PLoS One, 16, 1–12. https://doi.org/10.1371/journal.pone.0255537

Hara, P., Longjam, M., & Chattopadhyay, A. (2018). Stacking of mutant genes in the development of “purple tomato” rich in both lycopene and anthocyanin contents. Sci. Hortic., 239, 253–258. https://doi.org/10.1016/j.scienta.2018.05.039

Ikeda, H., Hiraoka, M., Shirasawa, K., Nishiyama, M., Kanahama, K., & Kanayama, Y. (2013). Analysis of a tomato introgression line, ILB-3, with increased resistance to brown spot. Sci. Hortic., 153, 103–108. https://doi.org/10.1016/j.scienta.2013.02.006

Kazvvozhi, T., Maether, P.M. (2003). The use of back propagating artificial neural networks in land cover classification. Int. J. Remote Sens. 24, 4907-4938. https://doi.org/10.1080/0143116030114851

Kohonen, T. (2014). MATLAB implementations and applications of the self-organizing feature map. Unigraf Oy, Helsinki, Finland.

Laviola, B. G., Silva, D. A. D., Jahn, A. C. P., Rocha, R. B., Oliveira, R. J. B., Albrecht, J. C., … Rosado, T. B. (2014). Desempenho agronomico e ganho genetico pela seleção de pinho-manso em três regiões do Brasil. Pesquisa Agropecuária Brasileira, 49, 356–363. https://doi.org/10.1590/0100-204X2014000500005

Londoño-Giraldo, L. M., Baena-Pedraza, A. M., Martínez-Seidel, F., Corpas-Ignacio, R., & Taborda-Ocampo, G. (2020). Genetic wild tomato introgression, physiological, volatilometric, and sensorial profiles ratify rustic relatives of cherry tomato as ideal mating partners. Sci. Hortic., 277, 1–10. https://doi.org/10.1016/j.scienta.2019.107981

Londoño-Giraldo, L. M., Gonzalez, J., Baena, A. M., Tapasco, O., Corpas, E. J., & Taborda, G. (2020). Selection of promising crops of wild cherry-type tomatoes using physicochemical parameters and antioxidant contents. Braga, 79, 169–179. https://doi.org/10.1590/0100-204X202000279

Maciel, G. M., Freitas, J. A., Maluf, W. R., Cardoso, M. G., & Benites, F. R. G. (2000). Melhoramento genético e ganho genético híbridos do tipo italiano. Revista Caatinga, 28, 1–4. https://doi.org/10.1590/S0100-204X2000040000004
