Journal of Agriculture and Food Research 15 (2024) 100993

Contents lists available at ScienceDirect

Journal of Agriculture and Food Research

journal homepage: www.sciencedirect.com/journal/journal-of-agriculture-and-food-research

Selfing revealed potential for higher yield performance than backcrossing among tomato segregating populations of *Solanum lycopersicum* × *S. pimpinellifolium* crosses under tropical humid climate

Chikezie Onuora Ene [a,b,*], Wosene Gebreselassie Abtew [a], Happiness Ogba Oselebe [c], Uchechukwu Paschal Chukwudi [e], Emeka Chibuzor Okechukwu [d,e], Friday Ugadu Ozi [c,f], Temesgen Matiwos Menamo [a], Chibueze Kelechi Ene [g], Agatha Ifeoma Atugwu [h]

[a] *Jimma University, Department of Horticulture and Plant Sciences, Jimma, Ethiopia*
[b] *Alex Ekwueme Federal University Ndufu-Alike, Department of Agriculture, Abakaliki, Nigeria*
[c] *Ebonyi State University, Department of Crop Production and Landscape Management, Abakaliki, Nigeria*
[d] *University of Worcester, School of Science and the Environment, Henwick Grove, Worcester, WR2 6 AJ, United Kingdom*
[e] *University of Nigeria, Department of Crop Science, Nsukka, Enugu State, Nigeria*
[f] *University of Wolverhampton, Department of Data Science and Artificial Intelligence, WV1 1LY, United Kingdom*
[g] *University of Potsdam, Institute of Biochemistry and Biology, 14476, Golm, Germany*
[h] *Natural Science Unit, School of General Studies, University of Nigeria, Nsukka, Nigeria*

ARTICLE INFO

Keywords:
Solanum pimpinellifolium
Genotype by yield*trait biplot
Humid environment
Multi-traits
Tomato population improvement
Superiority index

ABSTRACT

The objectives of this study were to assess and identify new source of phenotypic variability among F_3 and BC_1F_2 tomato populations, and apply genotype by yield*trait (GYT) biplots for population and line selection based on multiple traits. Four diverse cultivated parents ('CLN2498D' [D] and 'CLN2417H' [H] from Ethiopia; 'UC Dan INDIA' [U] and 'Tima' [T] from Nigeria), and wild parent 'LA2093' [W] were used to generate 276 potential breeding lines. The lines were categorized into eight populations ('pop_1_W/H1', 'pop_2_W/H2', 'pop_3_W/D1', 'pop_4_W/D2', 'pop_5_W/T1', 'pop_6_W/T2', 'pop_7_W/U1', and 'pop_8_W/U2'), and evaluated twice in the field using 19 × 15 alpha-lattice design with two replicates. Significant differences were observed among lines and populations for all yield enhancing traits. 'Pop_1_W/H1', 'pop_4_W/D2' and 'pop_6_W/T2' expressed the highest genetic divergence for plant height, number of leaves, total flower and fruit number, and fruit weight. GYT biplots revealed that all yield*trait interactions had a positive correlation with each other. F_3 populations, 'pop_5_W/T1' and 'pop_1_W/H1' exhibited the best performance for majority of the yield*trait combinations. Hierarchical clustering on principal components (HCPC) revealed overlapping lines (70.58% of Cluster D lines) and (54.05% of Cluster U lines) from the two F_3 populations. In BC_1F_2 population, 32.35% of the 34 original lines of Cluster D and 48.48% of Cluster T lines overlapped between Clusters D and T, while 18.18% of Cluster T lines and 8.82% of Cluster H lines were transgressive between Clusters T and H. Transgressive segregants '0210U1', '0211U1', and '0171T1' of selfed population using multivariate analysis were believed to represent potential sources of novel genetic variation for future tomato breeding.

1. Introduction

Tomato (*Solanum lycopersicum*) is the world's second most important vegetable crop after potato [1]. Its global production has increased by over 37% between 2002 and 2021 [2]. In 2021, 189.1 million metric tons of tomato was produced which was valued at about US $165.72 billion [2]. The increase in tomato fruit production hinges on the use of improved and high yielding cultivars [3].

Despite the increase in global tomato production, there is still a deficit in the global demand. Asfaw [4] observed a supply gap in tomato fruit demand. Atugwu et al. [5] attributed this yield deficit to environmental limitations placed on the most available high premium cultivars by the humid conditions. Tomato leaf size, fruit yields and fruit quality are reduced drastically at lower vapor pressure deficit or higher

* Corresponding author. Jimma University, Department of Horticulture and Plant Sciences, Jimma, Ethiopia.
E-mail address: enechike17@gmail.com (C.O. Ene).

https://doi.org/10.1016/j.jafr.2024.100993
Received 26 July 2023; Received in revised form 20 December 2023; Accepted 14 January 2024
Available online 19 January 2024

humidity. There is also significant increase in flower and fruit abortion, high disease occurrence and rapid fruit decay both in field and after harvest under high humidity conditions [6].

The environmental limitations imposed on tomato plants by high humidity leading to yield loss, reduced productivity and susceptibility to biotic and abiotic stressors have gained the attention of the stakeholders in tomato research and production. For instance, Xu [7] reported significant yield loss in tomato grown under high air humidity, which was suspected to be a result of lower photosynthesis-related activities and lower photosynthetic capacity evident under such growing condition compared to tomato exposed to low humidity. Alam et al. [8] observed extreme sensitivity of tomato to hot as well as humid environments, which affected yield and other physiological processes negatively. These studies attest to the need to address the environmental limitations imposed on tomato lines.

In tackling new challenges, breeders have relied on wild relatives of plant for novel gene donor. Among the wild relatives of tomato, *Solanum pimpinellifolium* stands out as a good gene donor. This is due to its large genetic variability, high compatibility, excellent growth architecture, good fruit organoleptic qualities, prolonged fruit shelf life, profuse flowering and fruiting, and high adaptability to humid environment [3, 9–16].

S. pimpinellifolium played a vital role in the development of heat-tolerant tomato varieties in Taiwan [17]. In Nigeria, it showed better responses to high humidity compared to the cultivated elite tomatoes [18]. Among many benefits of gene introgression from *S. pimpinellifolium* is the increased fruit number, which is a major contributory attribute to yield, in addition to resilience to high humidity [3]. These inherent benefits in *S. pimpinellifolium* is the motivation in selecting one of its lines 'LA2093' to improve tomato for increased yield and adaptability to the humid tropical climes.

Evaluation of different population means performance of derived progenies at early segregating generations as an approach to improve most economic traits especially yield and fruit quality traits for selection have recently been reported in tomato [19,20]. Selection of any of such populations for improvement would be dependent on performance of the populations across the multiple traits.

Genotype selection or population improvement based on multiplex traits raises a lot of concern in plant breeding. Genotype by yield*trait (GYT) biplot analysis is a novel tool with high efficiency compared to other statistical tools [21]. The GYT biplot helps to position genotypes or populations in hierarchy of performance based on their levels in combining yield with other traits other than yield, and expresses traits profiles of the genotypes or populations, which shows their strength and weaknesses. This is as opposed to the existing methods which only concentrate efforts on genotype evaluation by their levels in individual traits [22].

The objectives of this research are; (i) to assess the phenotypic variation in yield enhancing traits in tomato using segregating lines to identify novel sources of improved performance, and (ii) select population (s) among F_3 and BC_1F_2 populations, based on yield and other target traits using GYT biplots analyses in a humid environment.

2. Materials and methods

2.1. Experimental sites description

The present study was conducted at research field of the Department of Horticulture and Plant Sciences, Jimma University, Ethiopia during the 2020/2021 and 2021/2022 cropping seasons. Jimma is classified as warm to cold-environment locally known as "Weyna Dega", suitable for agriculture, with high degree of air humidity, located on latitude 07°4′N, longitude 36°50′E and altitude 1710 m above the sea level in the south west, Oromia region of Ethiopia [23]. The monthly weather conditions of Jimma during the experiment seasons are presented in Figs. 1 and 2. These figures show the mean monthly rainfall (amount), rain days, relative humidity and temperatures (minimum and maximum) in Jimma during the experimental seasons (2020/2021 and 2021/2022). There were variations in the minimum and maximum temperatures, amount of rainfall, rain days and relative humidity in the different years. In first year and second year, the highest rainfall occurred in November with 140.4 mm and 128.0 mm, followed by April with 118.0 mm and 110.3 mm, respectively. The amount of rainfall was lowest in January (11.3 mm and 13.3 mm) in the two years. The highest number of rainy days was 25 in April in the first season while November (23) took the lead in the second experiment.

The highest relative humidity recorded was in February (80%) followed by January (78%), April (77%) and March (75%) all in the second year, whereas in the first year, January and April shared the highest (74%) relative humidity. November in both planting years showed the least relative humidity recorded. During the two experimental years the maximum temperature range of 26 °C–28.7 °C and minimum temperature range of 9 °C–12.8 °C were recorded. In the first year, February and March recorded highest maximum (28.6 °C) and minimum (12.8 °C) temperatures while in the second year, November and April showed highest maximum (28.7 °C) and minimum (10.7 °C) temperatures, although temperature (maximum and minimum) data in the two years seemed statistically similar across the months. Jimma is mainly covered with black, gray and red colored plastic clay soils [24].

2.2. Genetic materials

A total of 276 potential breeding lines generated from four parental lines - elite parents of cultivated tomato (*S. lycopersicum*) and wild parent 'LA2093' of *S. pimpinellifolium* from California's C.M. Rick Tomato Genetic Resource Center, USA were used for this trial. The

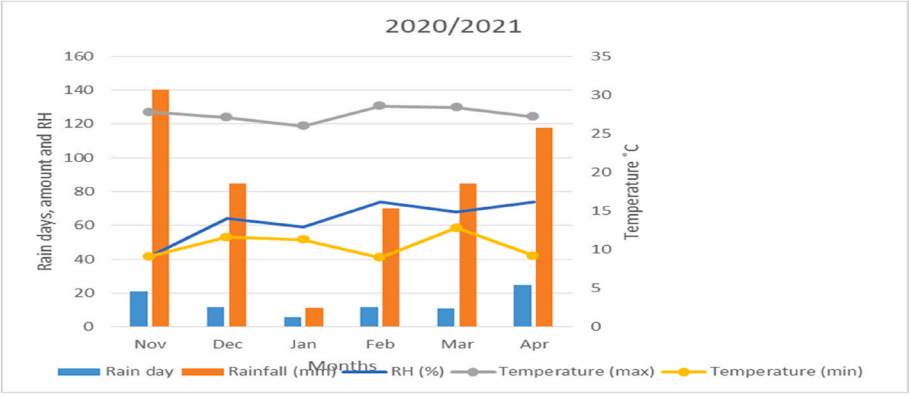

Fig. 1. Monthly weather conditions of Jimma during the 2020/2021 experiment season.

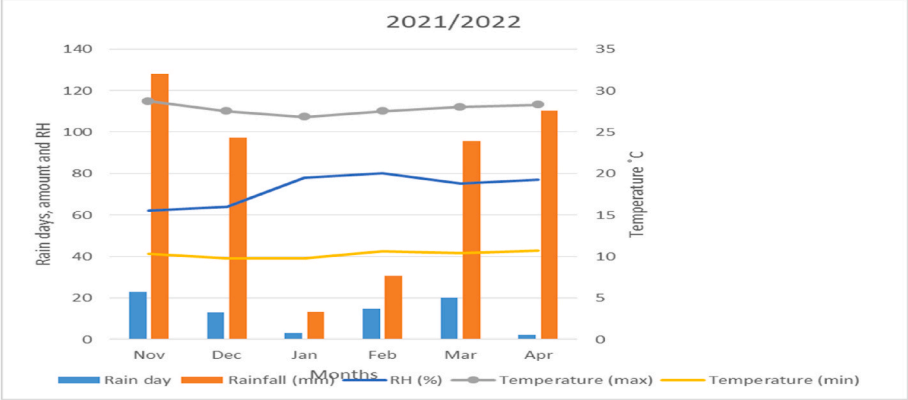

Fig. 2. Monthly weather conditions of Jimma during the 2021/2022 experiment season.

domesticated cultivars were morphologically diverse in the traits studied. The four elite lines - 'CLN2498D' (D) and 'CLN2714H' (H) from Ethiopia and 'UC Dan INDIA' (U) and 'Tima' (T) from Nigeria were used.

'CLN2498D' and 'UC Dan INDIA' were superior performing parents based on fruit yield and shape/size of fruits (round and large) but had less number of fruits as well as being susceptive of the humid condition of the growing environment. On the other hand, 'CLN2714H' and 'Tima' were poor performers in terms of fruit number, fruit yield, but had elongated medium fruit size, and longer floral and fruit phenology. The wild parent (W) 'LA2093' was commonly used as pollen donor to the four cultivated materials (females) using a bi-parental mating design by hand pollination according to Ozores-Hampton [25]. According to Wang et al. [3] 'LA2093' has prolonged shelf life, an early flowering and maturing variety, quantitative disease resistant and relatively resilient to humid environmental condition. It possesses enormous genetic potentials, having desirable alleles for breeding choice tomato cultivars with appreciable economic qualitative and quantitative traits performance.

Four F_1 crosses, *viz.*, 'H × W', 'T × W', 'D × W', and 'U × W' were obtained. Then two different crosses were done. Firstly, 40 F_1 plants from each cross were crossed back to the recurrent parents to obtain between 35–40 BCE_1 plants depending on each of the four crosses. The 35–40 BCE_1 plants were also selfed to obtain between 33–36 BC_1F_2 plants, and secondly, 40 F_1 plants were selfed to produce between 35 and 39 F_2 plants depending on each of the four crosses. The 35–39 F_2 plants were selfed to produce between 35 and 37 F_3 plants. A total of eight populations were developed: four F_3 populations and four BC_1F_2 populations. The detail of the crosses carried out and following populations were coded as presented in Table 1.

Table 1
Population naming description.

No.	Code	Number of lines	Description/pedigree
1.	pop_1_W/H1 (F_3 Cluster H)	35	F_3 population developed by selfing F_2 hybrids of a cross between 'CLN2714H' (H) and 'LA2093' (W)
2.	pop_2_W/H2 (BC_1F_2 Cluster H)	34	BC_1F_2 population developed by selfing backcross to the recurrent parent (BC_1) from a cross between 'CLN2714H' (H) and 'LA2093' (W); H = recurrent parent
3.	pop_3_W/D1 (F_3 Cluster D)	34	F_3 population developed by selfing F_2 hybrids of a cross between 'CLN2498D' (D) and 'LA2093' (W)
4.	pop_4_W/D2 (BC_1F_2 Cluster D)	34	BC_1F_2 population developed by selfing backcross to the recurrent parent (BC_1) from a cross between 'CLN2498D' (D) and 'LA2093' (W); D = recurrent parent
5.	pop_5_W/T1 (F_3 Cluster T)	33	F_3 population developed by selfing F_2 hybrids of a cross between 'Tima' (T) and 'LA2093' (W)
6.	pop_6_W/T2 (BC_1F_2 Cluster T)	33	BC_1F_2 population developed by selfing backcross to the recurrent parent (BC_1) from a cross between 'Tima' (T) and 'LA2093' (W); T = recurrent parent
7.	pop_7_W/U1 (F_3 Cluster U)	37	F_3 population developed by selfing F_2 hybrids of a cross between 'UC Dan INDIA' (U) and 'LA2093' (W)
8.	pop_8_W/U2 (BC_1F_2 Cluster U)	36	BC_1F_2 population developed by selfing backcross to the recurrent parent (BC_1) from a cross between 'UC Dan INDIA' (U) and 'LA2093' (W); U = recurrent parent
Total		276	

2.3. Field management and experimental design

The plant materials were evaluated using an alpha lattice experimental design with two replicates. Each replicate contained 19 blocks whereas each block contained 15 plots. The tomato seeds were first planted in plastic seed trays filled with sterilized top soil mixed with well cured poultry manures and river sand in the ratio of 3: 2: 1, respectively by volume with the use of a head pan. At 24 days after seedling emergence with the appearance of 4–5 true leaves, uniform vigorous seedlings were transplanted to the field.

Each plot occupied by a particular genotype of each cross whether $F_{3:4}$ or $BC_1F_{2:3}$ consisted of three rows of plant (30 plants/plot) in an area of 1.5 m × 5 m with 0.5 m × 0.5 m inter-intra row spacing. A distance of 1 m and 0.5 m alleys were ensured between blocks and plots, respectively. Poultry droppings at the rate of 10 metric tons per hectare were worked into the soil within each replicate 16 days before seedlings were transplanted. Cultural practices such as weed control, irrigation, fertilizer (DAP-Di Ammonium Phosphate), fungicide (Ridomil-Mancozeb and Metalaxyl-M), staking, pruning, and insecticide (Karate-Lambda-Cyhalothrin 5% EC) were applied as suggested by Osei et al. [26].

2.4. Data collection

Data were taken on 10 plants of each genotype in each replicate all from inner plants of each plot eluding border plants. The following morphological traits were measured and recorded at 9 week after transplanting during which the genotypes were expected to have completed vegetative process as recommended by AVRDC guideline [27].

Plant height (PH cm) was measured from the plant base to the shoot tip. Number of leaves (NL), primary branches (NBp), secondary branches (NBs), and nodes (NN) were counted. Phenological traits included; number of days to first anthesis (DFA), 50% anthesis (D50A), first fruit emergence (DFFE), 50% fruit set (D50FS), and first fruit ripening (DFFR) were taken from the day of transplantation of seedlings. Fruit shelf-life traits included: number of days to first fruit spoilage

(D1stFSp) and 100% fruit spoilage (D100FSp) stored under room temperature according to Arah et al. [28].

The room temperature ranged from 18 °C to 23 °C while the relative humidity was from 84% to 91% throughout the shelf life duration. Total number of flowers per plant (TNFlPP) and total number of fruits per plant (TNFrPP) were counted. Fruits were cut crosswise to count number of locules per fruit (NLPF). The total number of mature fruits showing ripening initiation at second harvest was weighed with electronic weighing balance and the fruit weight per plant (FWPP g) recorded. The total fruit yield per hectare (TFYPH t/h) was estimated as described by Dinssa et al. [27].

2.5. Statistical analysis

Before computing the analysis of variance, test for homogeneity of residual variances was carried out using F test where larger variance was divided by the smaller variance between the two seasons and all showed less than three which is the threshold, and suggested the two sets of data could be combined as recommended by Gomez and Gomez [29]. Each population data along parents was subjected to two-way analysis of variance (ANOVA) using R software to analyze variance components [30].

Descriptive statistics and genetic parameters were done using the package "variability" in R as recommended by Raj et al. [31]. The descriptive statistics were means, and ranges (minimum and maximum), while the genetic parameters included genetic advance as percentage of mean (GAM), genotypic coefficient of variation (GCV), phenotypic coefficient of variation (PCV); broad sense heritability (h_{bs}) which was estimated as the ratio of genotypic variance (GV) to phenotypic variance (PV). The GCV and PCV values were categorized as high (>20%), medium (10–20%) and low (<10%) based on the recommendation of Burton and DeVane [32]. The expected GAM (%) on 5% selection intensity were ranked as low (1–10%), intermediate (10–20%) and high (>20%) whereas broad sense heritability was categorized as low (0–30%), intermediate (30–60%) and high (>60%) according to Singh and Choudhary [33] and Khan et al. [34]. Comparison among populations was based on these classified genetic parameters: GCV, PCV, h_{bs} and GAM using graphs.

Data were further analyzed by the 'genotype by yield*traits (GYT)' biplots for the 'Tester Vector' view showing relationship among yield-trait combinations, the 'Average Tester Coordination' view ranking the genotypes based on their overall superiority and their strengths and weaknesses, and 'which wins where' view or the polygon view highlighting genotypes with outstanding profiles based on the genotype by yield*trait combined data, an option of GGE biplot software version 6.3 [35]. We adopted the principle of GYT biplot analysis on the various populations with 'each population mean' representing genotype of the GYT biplot analysis.

From the original data (Supplemental Table S1), a GYT table was derived (Supplemental Table S2), in which each column was a yield*-trait combination. The method indicates that each trait should either be multiplied or divided by total fruit yield according to the objectives of breeding. To develop the GYT table, traits with favorable or desirable increased values (such as PH, NL, NBp, NBs, NN, TNFlPP, TNFrPP, FWPP, D1stFSp, D100FSp, and NLPF) were multiplied with yield using the multiplication operator ("*") whereas traits with favorable decreased values (the phenological traits) were used to divide yield using the division operator ("/").

By this method, in the GYT table a larger value is always more desirable. A mean superiority index (MSI) which involved all yield*trait combinations was also calculated based on the standardized GYT table as described by Yan and Frégeau-Reid [21] using the formula as follows:

$$P_{ij} = \frac{T_{ij} - T_j}{S_j}$$

where,

Pij is the standardized value of genotype *i* for yield-trait combination *j* in the standardized table, Tij is the original value of genotype *i* for yield-trait combination *j* in the GYT table (Supplemental Table S2), *Tj* is the mean across genotypes for yield-trait combination *j*, and *Sj* is the standard deviation for yield-trait combination *j*. The biplots of GYT were based on singular value decomposition of trait-standardized data ("Scaling = 1, Centering = 2") and trait-focused singular value partition ("SVP = 2") according to Yan and Tinker [36], for all the views employed in the present trial except the 'Average Tester Coordination' view which showed SVP = 1. Traits were regarded as 'tester' when using 'relation among testers' option. The 'Tester Vector' view of the GYT biplots showed associations among the yield-trait combinations. The 'Average Tester Coordination' view of the GYT biplots ranked the populations based on their overall superiority and their strengths and weaknesses. The 'which-won-where' view of the GYT biplots was used to highlight population with outstanding profiles.

Hierarchical clustering on principal components (HCPC) was done to group the genotypes based on the measured traits, and the results were visualized using the *fviz_cluster* functions of the R package *"factoextra"* for factor map [37].

3. Results

3.1. Genetic variability, heritability and genetic advance for quantitative traits

Despite the variations recorded in some weather elements for the two years, the test for homogeneity carried out on the tomato traits data suggested the performance of a combined analysis of variance. The result showed that genotypes within each population differed significantly for total fruit yield and other traits studied ($P < 0.001$). The descriptive and genetic parameters including h_{bs} and GAM of the traits in each population are summarized in Supplemental Tables S4a & S4b to S7a & S7b.

Comparing the parental lines, 'CLN2498D' was the highest for NLPF (7.22), FWPP (3093.2 g) and yield (64.03 t/h), followed by 'UC Dan INDIA' (6.48, 2955.1 g and 61.17 t/h, respectively) while the wild parent, 'LA2093' displayed the highest values for PH (160.06 cm), NL (191.48), NBp (12.53), NBs (45.72), NN (39.51), TNFlPP (511.78) and TNFrPP (496.25), highest days to observe the fruit firmness and expressed the least days to observed all the floral and fruit phenological traits. Variation of the populations from the better performing parent was also shown in this study. For instance, 'pop_1_W/H1' showed higher maximum values for PH (176.29 cm), TNFlPP (568.66), TNFrPP (534.15) and NLPF (7.35); 'pop_7_W/U1' for NL (195.96), NBp (14.53), NBs (47.81), and NN (43.34); 'pop_8_W/U2' for D1stFSp (27.24 days) and D100FSp (37.23 days) than those recorded in the better parents depending on trait.

'Pop_5_W/T1' had the highest maximum value among the 8 populations for FWPP (2292.55 g) and fruit yield (46.15 t/h), the least maximum values for DFFE (28.23 days), and DFFR (48.82 days); 'pop_7_W/U1' for DFA (15.93 days), D50A (20.04 days) and D50FS (36.81 days) although the values were poorer than the better performing parents. Moreover, the mean values showed similar results trend with the maximum range descriptions except that among the populations, least D50FS (34.89 days) was recorded in 'pop_5_W/T1'.

All the 8 populations displayed low (<10%) GCV and PCV for the phenological traits: DFA, D50A, DFFE, D50FS and DFFR. High (>20%) GCV and PCV were shown in all the populations for PH. For NL, majority of the populations expressed medium to high GCV and PCV with the exception of 'pop_4_W/D2' and 'pop_8_W/U2'. Intermediate (10–20%) or high (>20%) GCV and PCV were recorded for FWPP and NLPF in all populations except 'pop_5_W/T1' which showed low values for FWPP. All the populations displayed medium or high GCV and PCV values for NBp except 'pop_8_W/U2', whereas for NBs all the populations were low

apart from 'pop_1_W/H1', 'pop_5_W/T1' and 'pop_6_W/T2'.

'Pop_3_W/D1' and 'pop_7_W/U1' showed high GCV and PCV for NN; leaving the rest populations intermediate, whereas 'pop_6_W/T2' had low values. 'Pop_3_W/D1', 'pop_4_W/D2', 'pop_7_W/U1' showed intermediate GCV and PCV values for TNFlPP and TNFrPP while 'pop_5_W/T1' and 'pop_6_W/T2' could display similar category of GCV and PCV only for TNFrPP. For the two traits which tested fruit shelf life after harvest, D1stFSp and D100FSp, all the populations expressed medium or high GCV and PCV values with the exclusion of 'pop_3_W/D1' and 'pop_4_W/D2', although 'pop_7_W/U1' displayed medium GCV and PCV only for D100FSp. For fruit yield, 'pop_2_W/H2', 'pop_3_W/D1' and 'pop_7_W/U1' displayed moderate GCV and PCV values. The difference between the GCV and PCV values were negligible for all the traits in all populations.

All the populations exhibited very high magnitude of GV and PV for traits such as PH, NL, TNFlPP, TNFrPP, and FWPP. Among the populations, 'pop_1_W/H1' displayed in magnitude the highest GV and PV values for PH and NL; 'pop_4_W/D2' for TNFlPP and TNFrPP; and 'pop_6_W/T2' for FWPP. The variations that existed between the PV and GV for all the traits in all the populations were minimal.

High (>20%) GAM was displayed by all the populations for PH, NL, NBp, NN, and NLPF, except 'pop_4_W/D2' for NL (13.13%); 'pop_6_W/T2' for NN (16.20%) and 'pop_8_W/U2' for NL (13.18%) and NBp (15.42%) which showed intermediate values. For NBs, high GAM were expressed by 'pop_1_W/H1', 'pop_5_W/T1' and 'pop_6_W/T2'; intermediate by 'pop_2_W/H2', 'pop_3_W/D1' and 'pop_7_W/U1'. GAM values for DFA and D50A were low (<10%) in 'pop_1_W/H1', 'pop_4_W/D2' and 'pop_6_W/T2' and moderate (10–20%) in 'pop_3_W/D1' and 'pop_8_W/U2'. However, 'pop_2_W/H2' displayed moderate value for DFA whereas 'pop_5_W/T1' and 'pop_7_W/U1' exhibited similar class of GAM for D50A.

The expression of high GAM was observed in 'pop_3_W/D1', 'pop_4_W/D2', and 'pop_7_W/U1' for TNFlPP and TNFrPP leaving the rest with medium values for similar traits. All the populations showed low (<10%) GAM for traits such as: DFFE, D50FS, and DFFR except 'pop_6_W/T2' which displayed a weak intermediate value for DFFE. Apart from 'pop_5_W/T1' which expressed moderate GAM for FWPP, the rest populations showed high values with 'pop_3_W/D1' being the highest. For D1stFSp and D100FSp, high GAM were noted in 'pop_1_W/H1', 'pop_2_W/H2', 'pop_5_W/T1' and 'pop_6_W/T2'; intermediate values were observed in 'pop_3_W/D1' and 'pop_4_W/D2'; while 'pop_7_W/U1' displayed high GAM for D100FSp and moderate for D1stFSp. All the populations evaluated showed medium GAM for total fruit yield except 'pop_2_W/H2', 'pop_3_W/D1' and 'pop_7_W/U1' which displayed high values for the same trait.

High (>60%) broad sense heritability were found in all the populations for virtually all the traits studied. However, moderate (30–60%) heritability estimates were consistently observed for DFFR in all the populations except 'pop_6_W/T2' (63.20%) and 'pop_5_W/T1' (26.25%) which showed high and low (<30%) values, respectively for similar trait. 'Pop_4_W/D2' expressed intermediate heritability for DFFE (52.03%) and low value for D50FS (25.97%) leaving the rest populations with high values for the same traits. Moderate heritability estimate was recorded in 'pop_6_W/T2' for D50A (50.50%) among all the populations. Figs. 3–7 present graphs which compared the 8 populations for each quantitative trait and concentrated only on the categorized genetic parameters such as genotypic coefficient of variation, phenotypic coefficient of variation, heritability estimates in broad sense and genetic advance as a percentage of mean.

3.2. Genotype by yield*trait (GYT) biplots using different views

The idea behind the involvement of different views of the GYT biplot was to allow the data to be investigated from different angles for better understanding of the relationship between populations and yield*trait combinations, superior population(s) as well as the yield*trait profiles of

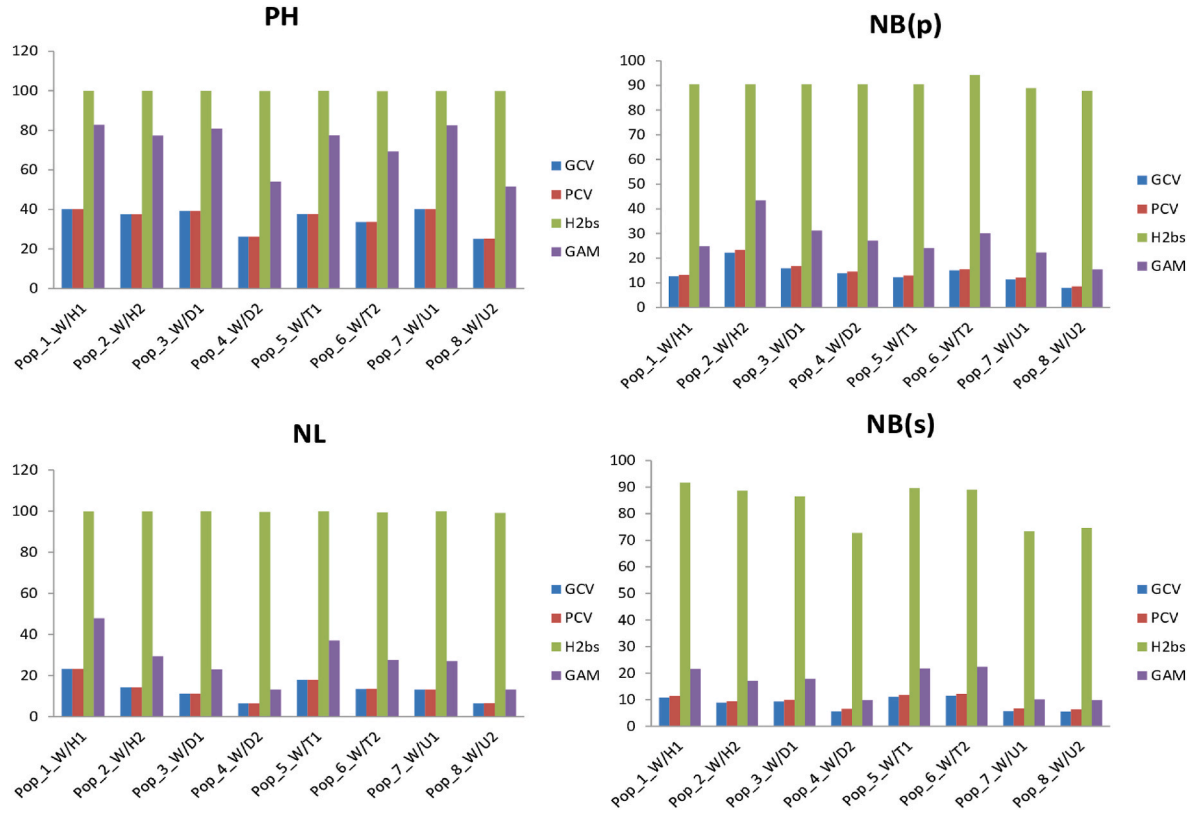

Fig. 3. Genetic variation, broad sense heritability and genetic advance among tomato populations for plant height (PH), number of primary branches (NBp), number of leaves (NL) and number of secondary branches (NBs).

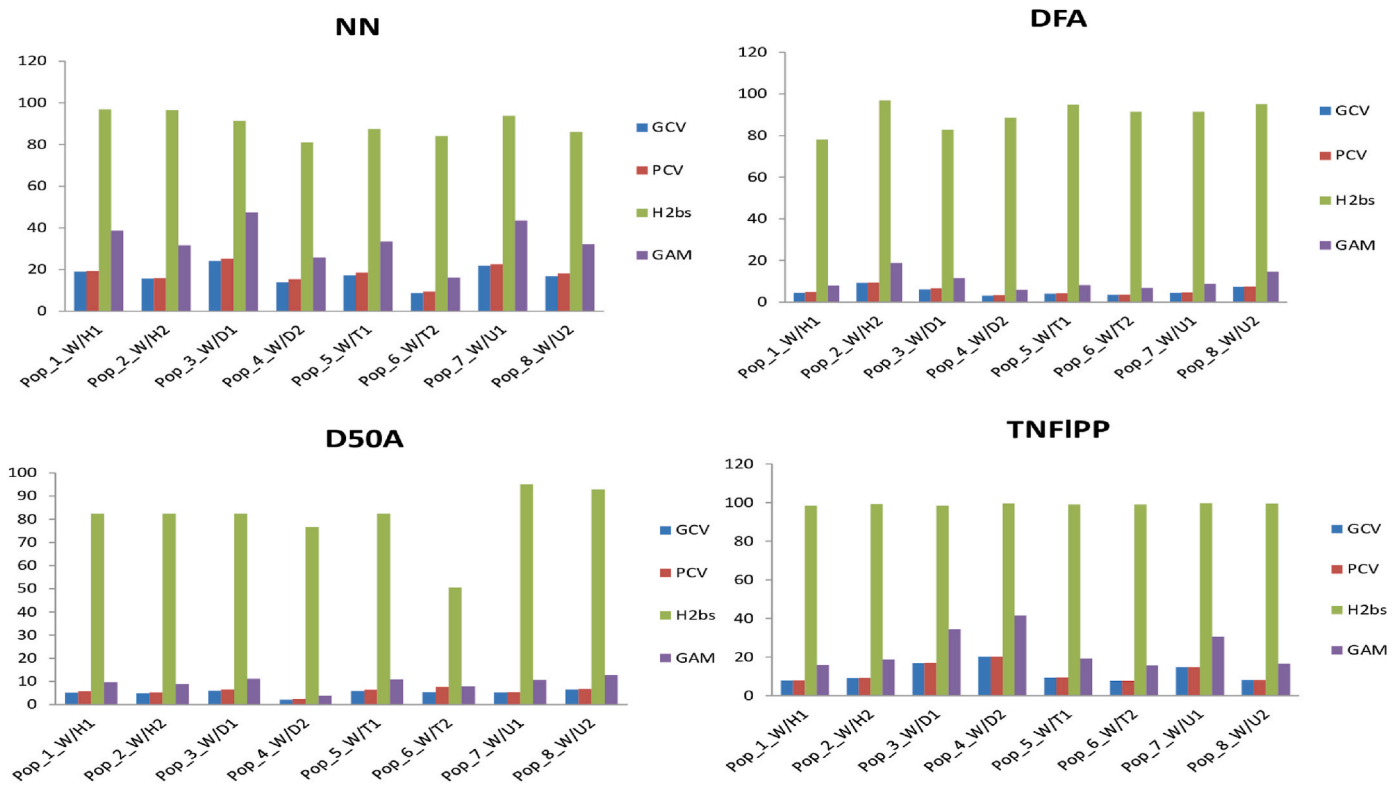

Fig. 4. Genetic variation, broad sense heritability and genetic advance among tomato populations for number of nodes (NN), days to first anthesis (DFA), days to 50% anthesis (D50A) and total number of flower per plant (TNFlPP).

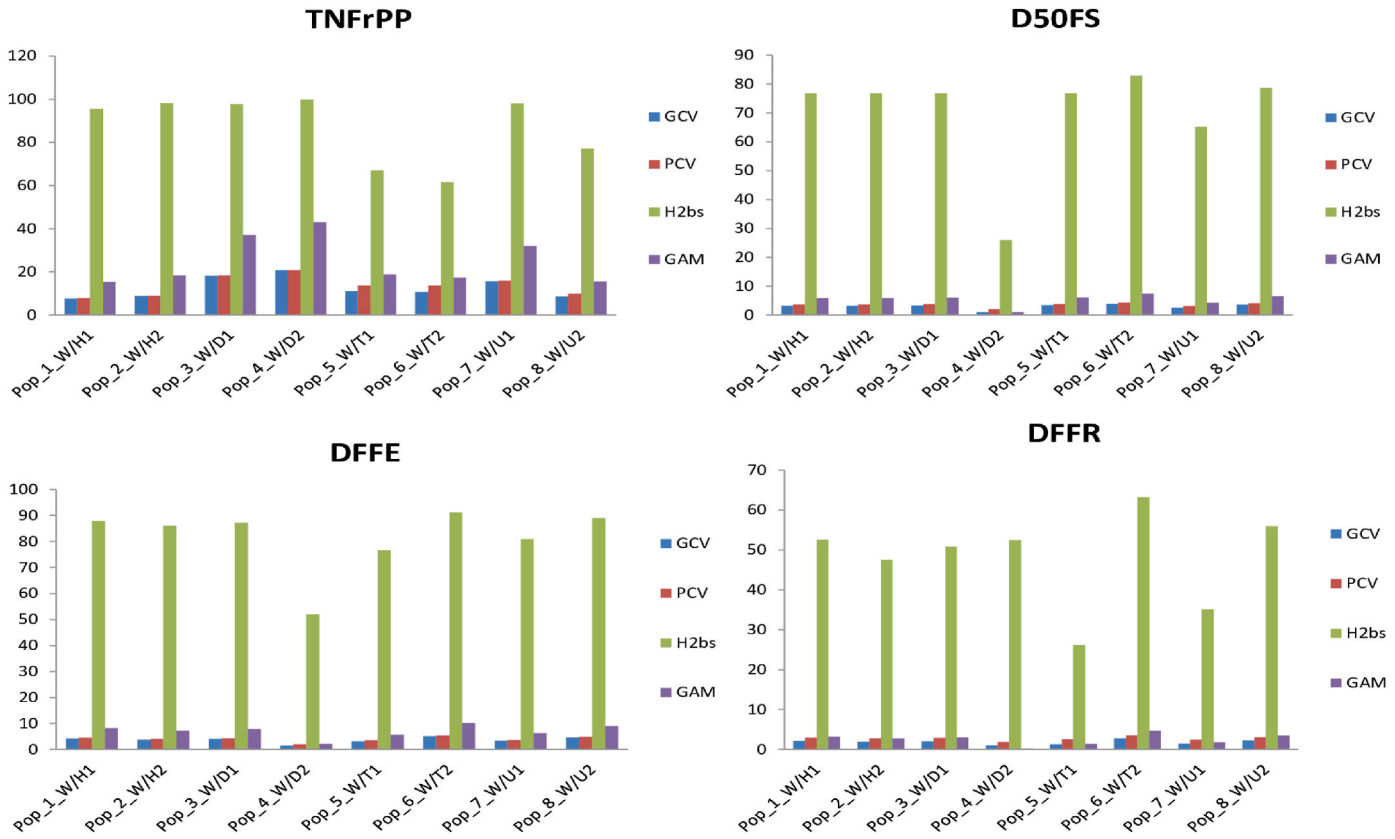

Fig. 5. Genetic variation, broad sense heritability and genetic advance among tomato populations for total number of fruit per plant (TNFrPP), days to 50% fruit set (D50FS), days to first fruit emergence (DFFE) and days to first fruit ripening (DFFR).

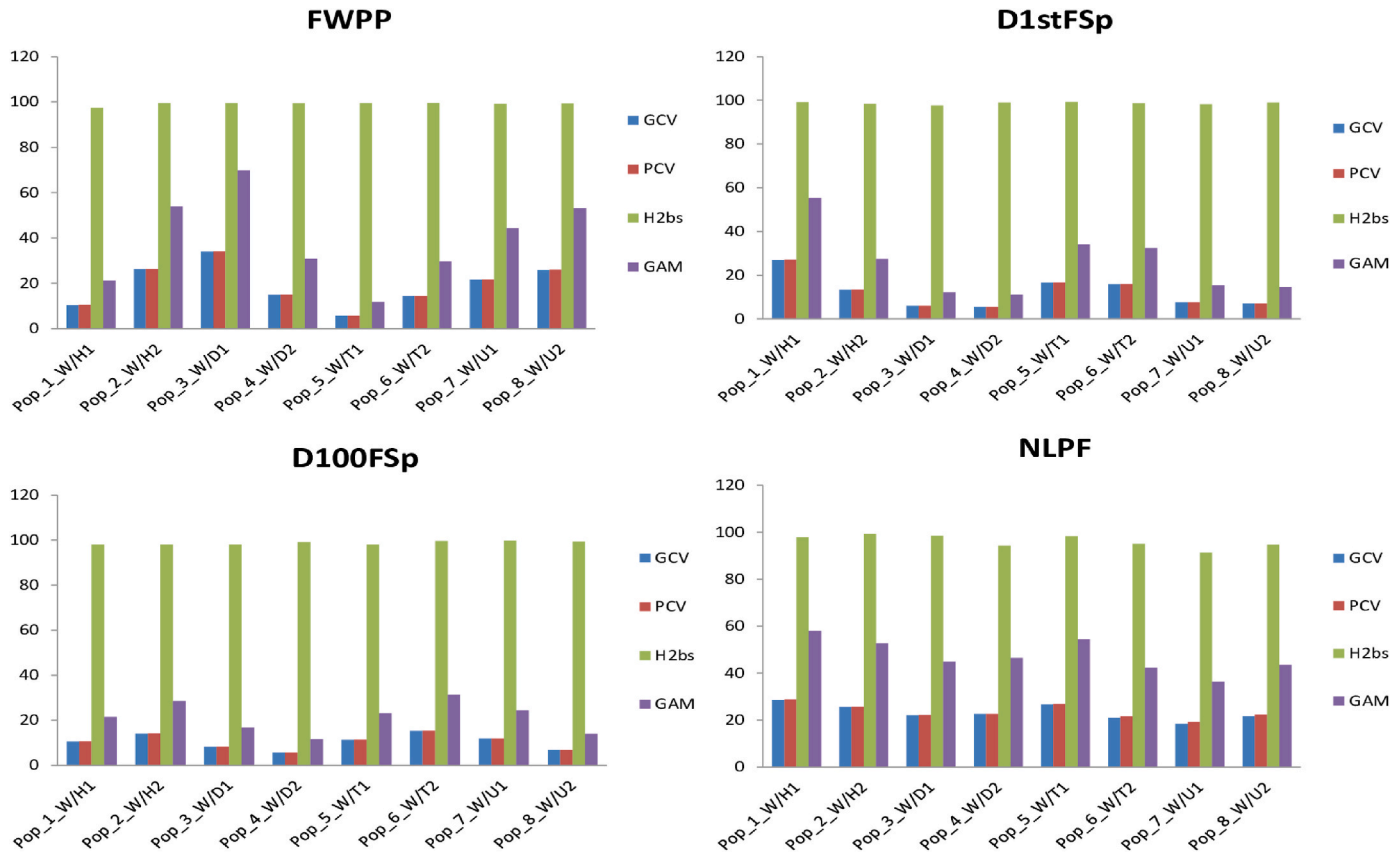

Fig. 6. Genetic variation, broad sense heritability and genetic advance among tomato populations for fruit weight per plant (FWPP), days to first fruit spoilage (D1stFSp), days to 100% fruit spoilage (D100FSp) and number of locules per fruit (NLPF).

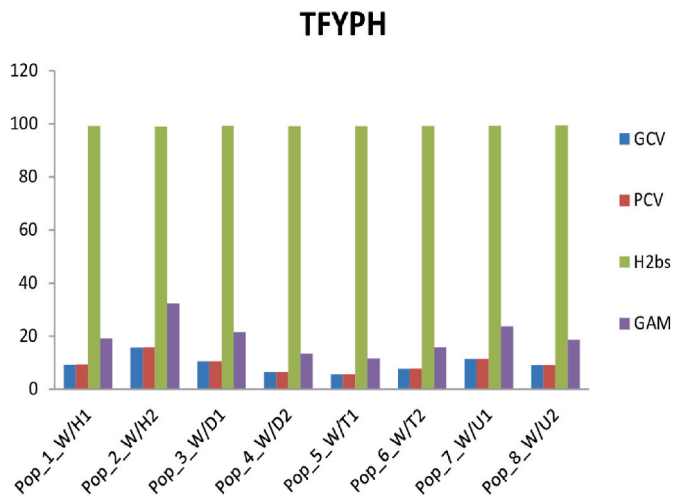

Fig. 7. Genetic variation, broad sense heritability and genetic advance among tomato populations for total fruit yield per hectare.

the populations. The total variation revealed by the PC1 and PC2 among the yield*traits combinations of these views/models was 97.5%.

*3.2.1. Relationships among yield*trait combinations using the Tester Vector view of the GYT biplot*

According to the Tester Vector view of the GYT biplot for the 8 populations, all yield*trait interactions had a positive correlation with each other as shown by the acute angles between their vectors (Fig. 8). This may be due to the involvement of yield as a component of all yield*trait combinations. This vital feature of GYT biplot makes it unique when compared to GT biplot as genotypes or populations are easily ranked graphically and meaningfully based on their yield*trait combinations.

The result showed that the magnitudes of angles among yield*D1stFSp, yield*D100FSp and the rest yield*trait combinations were higher although all yield*trait combinations indicated high positive correlation. Among all populations, 'pop_5_W/T1' and 'pop_1_W/H1' had the largest values for yield*PH, yield*NN, yield*TNFlPP and yield*TNFrPP, although they had high proximity with several other yield*trait combinations.

3.2.2. Superiority vs. "weaknesses and strengths" of genotypes using the Average Tester Coordination view of the GYT biplot

Fig. 9 represents the Average Tester Coordination view for the 8 tomato populations. This view of the GYT biplot showed the superiority ranking of the populations based on their yield*trait combinations. The single arrow line which passed through the biplot origin and the average yield*trait interactions is the average tester coordinate (ATC) whereas the small circle on the ATC showed the placement of the "average yield*trait combination," which was determined by the coordinates of all yield*trait combinations in the biplot. The double arrow blue line is the general population mean. It divided the plot into two parts which separated the populations based on their performance. The populations found on the side of the ATC arrow (right) showed better performance for the surrounding yield*traits combinations whereas those found at the opposite side (left) showed weak performance.

The best ranked population based on the yield*trait combinations was 'pop_5_W/T1' and was followed by 'pop_1_W/H1', 'pop_4_W/D2', and 'pop_6_W/T2'. In contrast, 'pop_2_W/H2', 'pop_7_W/U1', 'pop_3_W/D1', and 'pop_8_W/U2' were ranked the poorest based on their

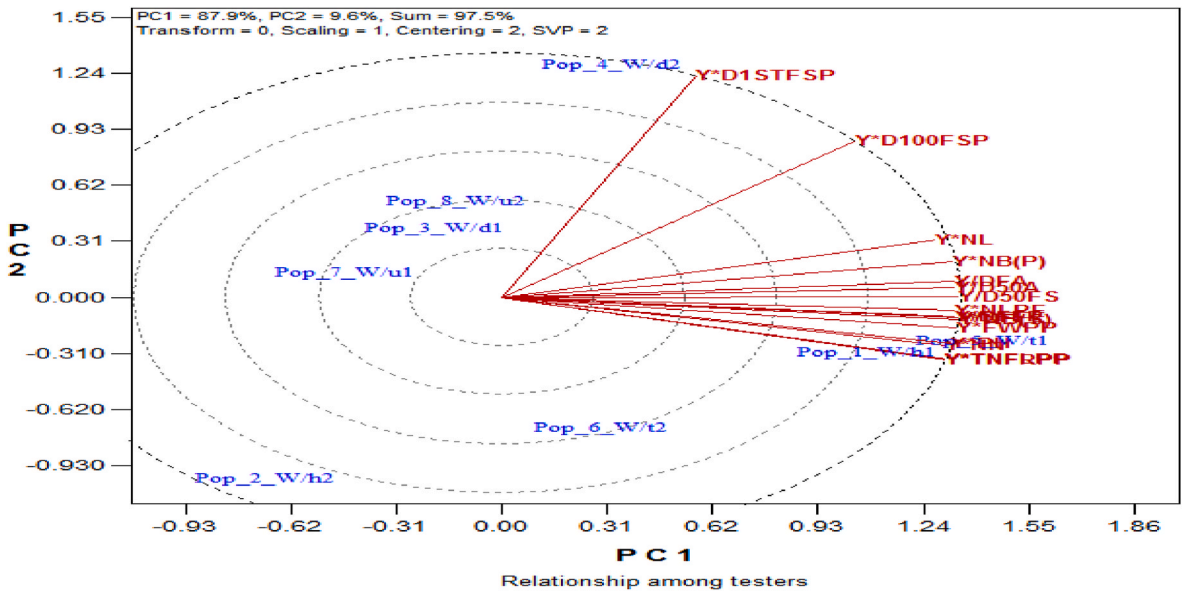

Fig. 8. The Tester Vector view showing relationship among testers/traits of genotype by yield*trait (GYT) biplot of 8 populations based on yield*traits combinations; PH (cm): Plant height, NL: Number of leaves, NB(p): Number of primary branches, NB(s): Number of secondary branches, NN: Number of nodes (at 9 week after transplanting), DFA: Days to first anthesis, D50A: Days to 50% anthesis, TNFlPP: Total number of flower per plant, TNFrPP: Total number of fruit per plant, FWPP (g): Fruit weight per plant, DFFE: Days to first fruit emergence, D50FS: Days to 50% fruit set, DFFR: Days to first fruit ripening, D1stFSp: Day to initial fruit spoilage after harvest, D100FSp: Days to 100% fruit spoilage after harvest, NLPF: Number of locule per fruit, Y (t/h): Total fruit yield per hectare, pop_1_W/H1 (F_3 of 'H × W'), pop_2_W/H2 (BC_1F_2 of 'H × W'), pop_3_W/D1 (F_3 of 'D × W'), pop_4_W/D2 (BC_1F_2 of 'D × W'), pop_5_W/T1 (F_3 of 'T × W'), pop_6_W/T2 (BC_1F_2 of 'T × W'), pop_7_W/U1 (F_3 of 'U × W'), pop_8_W/U2 (BC_1F_2 of 'U × W').

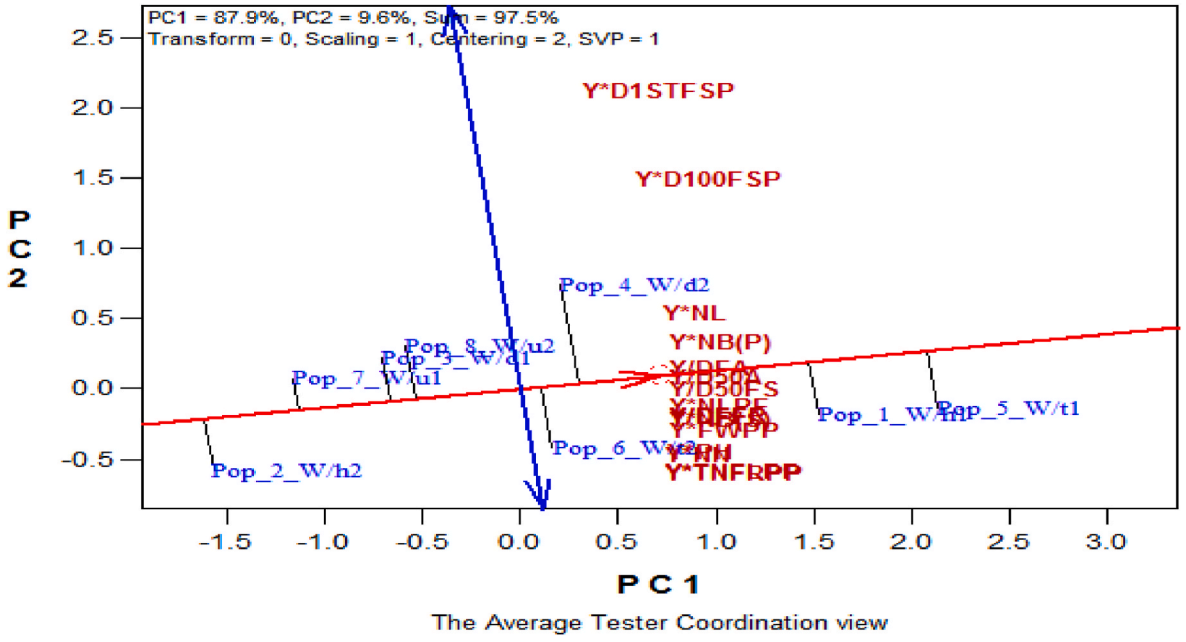

Fig. 9. The Average Tester Coordination view of genotype by yield*trait (GYT) biplot involving 8 populations based on multi-trait; PH (cm): Plant height, NL: Number of leaves, NB(p): Number of primary branches, NB(s): Number of secondary branches, NN: Number of nodes (at 9 week after transplanting), DFA: Days to first anthesis, D50A: Days to 50% anthesis, TNFlPP: Total number of flower per plant, TNFrPP: Total number of fruit per plant, FWPP (g): Fruit weight per plant, DFFE: Days to first fruit emergence, D50FS: Days to 50% fruit set, DFFR: Days to first fruit ripening, D1stFSp: Day to initial fruit spoilage after harvest, D100FSp: Days to 100% fruit spoilage after harvest, NLPF: Number of locule per fruit, Y (t/h): Total fruit yield per hectare, pop_1_W/H1 (F_3 of 'H × W'), pop_2_W/H2 (BC_1F_2 of 'H × W'), pop_3_W/D1 (F_3 of 'D × W'), pop_4_W/D2 (BC_1F_2 of 'D × W'), pop_5_W/T1 (F_3 of 'T × W'), pop_6_W/T2 (BC_1F_2 of 'T × W'), pop_7_W/U1 (F_3 of 'U × W'), pop_8_W/U2 (BC_1F_2 of 'U × W').

performance below the population mean for all the yield*trait combinations. Based on traits profiles of the populations, 'pop_5_W/T1', 'pop_1_W/H1', and 'pop_6_W/T2' appeared to be balanced for various yield*trait combinations; while 'pop_4_W/D2' was strong in the two shelf-life traits: D1stFSp and D100FSp in combination with yield but weak in TNFlPP, TNFrPP, PH, and NN. Despite populations general superiority or weakness stand, those placed below the ATC such as 'pop_5_W/T1', 'pop_1_W/H1', 'pop_6_W/T2', and 'pop_2_W/H2' tended

to have relatively good levels of the traits below the ATC while having poor levels of the traits above the ATC. The opposite is true for populations ('pop_7_W/U1', 'pop_3_W/D1', 'pop_8_W/U2' and 'pop_4_W/D2') placed above the ATC.

The weakness and strength of each population is presented in Supplemental Table S3. The best populations were 'pop_5_W/T1' and 'pop_1_W/H1' based on the magnitude of their mean values as expressed in the mean superiority index (MSI) table, in that order. Of these, 'pop_5_W/T1' did not have any negative value for all yield*traits combinations, while the second placed 'pop_1_W/H1' exhibited negative values for D1stFSp. Apart from these best performer populations mentioned which had positive MSI, the rest populations expressed negative values for MSI.

3.2.3. *Genotypes performance on yield*trait combinations using which-won-where view of the GYT biplot*

The result showed that 'pop_5_W/T1' accompanied by 'pop_1_W/H1' exhibited the best performance for a majority of the yield*trait combinations (Fig. 10). The only yield*trait combination left out was yield*D1stFSp which projected 'pop_4_W/D2' as the best performer. However, 'pop_4_W/D2', followed by 'pop_8_W/U2', and 'pop_5_W/T1' showed leading performance for yield*D100FSp which appeared on the radiate line. The result showed that these populations mentioned were the best in combining fruit yield with the other traits as observed in the present biplot.

3.3. *Clustering pattern analysis for the quantitative traits in F_3 and BC_1F_2 populations*

Clustering pattern analysis using hierarchical clustering on principal components (HCPC) classified the lines into four clusters in the two population types (Figs. 11 and 12). However, transgressive segregation was observed for the studied traits in both F_3 and BC_1F_2.

There was significant structuring among the lines observed in F_3 composite populations. Among F_3 populations (Fig. 11), clusters D and U which encompassed originally lines from crosses between 'CLN2498D × LA2093' and 'UC Dan INDIA × LA2093' contained transgressive or overlapping lines (70.58% of all D lines) and (54.05% of all U lines), respectively from the two populations. These two populations/clusters were characterized by less vigorous plants with lower performance in virtually all traits compared to clusters H and T (Supplemental Fig. S1, Supplemental Tables S4a, S5a, S6a, & S7a). Of the 37 lines of 'UC Dan INDIA × LA2093' cross (cluster U), two lines '0210U1' and '0211U1' were lone genotypes (outliers) whereas '0171T1' was the only lone genotype from 'Tima × LA2093' cross (cluster T 'pop_5_W/T1').

Similar trend was found between clusters H and T of 'CLN2417H × LA2093' and 'Tima × LA2093' crosses, respectively. 8.57% of all H lines and 30.30% of all T lines were found overlapping between clusters/populations H and T. Lines such as '0020H1' and '0035H1' of cluster H found their way expressly among lines of cluster T whereas '0144T1' of cluster T showed slight drifting into cluster H. The cluster description is found in Fig. 11.

The mean of each F_3 population plotted in a boxplot analysis displayed significant differences for all the studied traits (Supplemental Fig. S1). Cluster H 'pop_1_W/H1' had the highest mean values for PH, TNFlPP, and TNFrPP; while cluster T 'pop_5_W/T1' had the highest for NB_prim, NB_sec, NN, FWPP, and fruit yield, the least DFFE, DFFR, also taking the shortest days to 100% fruits shriveling. Number of leaves, D1stFSp, and D100FSp implicated cluster/population U 'pop_7_W/U1' as the cluster with the highest mean values while having the least mean values for DFA and D50A. In contrast, Clusters D 'pop_3_W/D1' and U consistently showed lower performance comparatively for traits such as PH, NB_prim, NB_sec, NN, TNFlPP, TNFrPP, including fruit yield.

With respect to BC_1F_2 populations (Fig. 12), lots of transgressive segregations were also observed. Three clusters, D, T and H showed profuse overlapping *viz.*, clusters D and T, as well as clusters T and H.

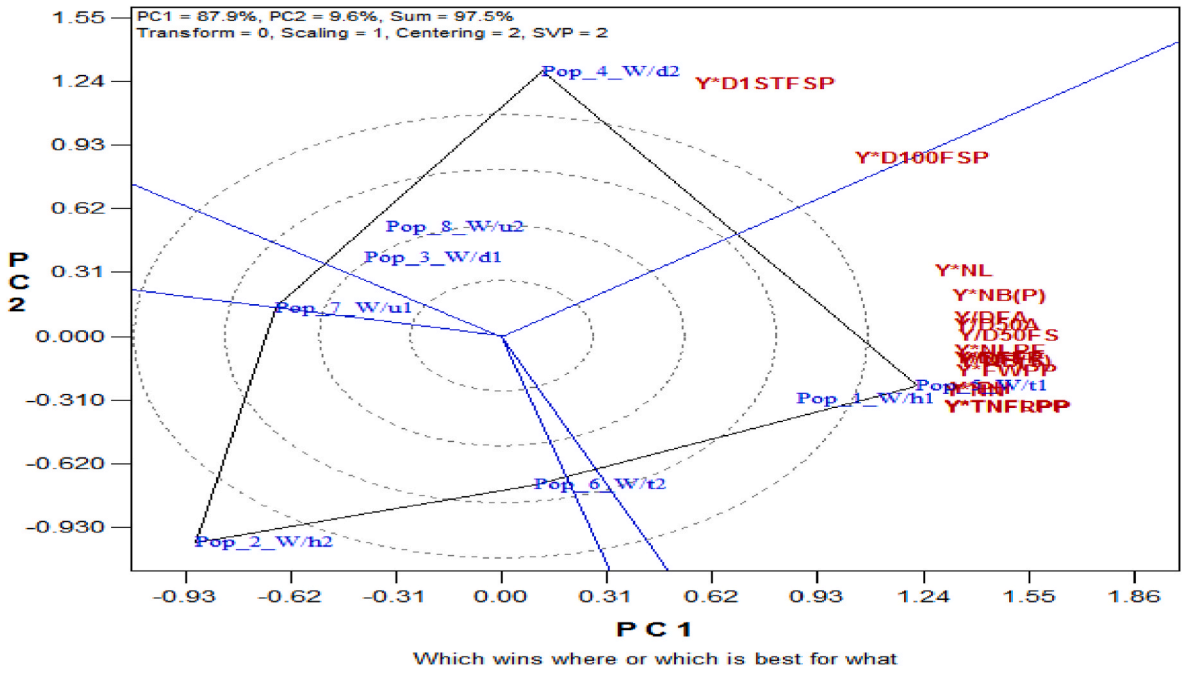

Fig. 10. The which-won-where view of genotype by yield*trait (GYT) biplot involving 8 populations based on multi-traits; PH (cm): Plant height, NL: Number of leaves, NB(p): Number of primary branches, NB(s): Number of secondary branches, NN: Number of nodes (at 9 week after transplanting), DFA: Days to first anthesis, D50A: Days to 50% anthesis, TNFlPP: Total number of flower per plant, TNFrPP: Total number of fruit per plant, FWPP (g): Fruit weight per plant, DFFE: Days to first fruit emergence, D50FS: Days to 50% fruit set, DFFR: Days to first fruit ripening, D1stFSp: Day to initial fruit spoilage after harvest, D100FSp: Days to 100% fruit spoilage after harvest, NLPF: Number of locule per fruit, Y (t/h): Total fruit yield per hectare, pop_1_W/H1 (F_3 of 'H × W'), pop_2_W/H2 (BC_1F_2 of 'H × W'), pop_3_W/D1 (F_3 of 'D × W'), pop_4_W/D2 (BC_1F_2 of 'D × W'), pop_5_W/T1 (F_3 of 'T × W'), pop_6_W/T2 (BC_1F_2 of 'T × W'), pop_7_W/U1 (F_3 of 'U × W'), pop_8_W/U2 (BC_1F_2 of 'U × W').

Fig. 11. Factor map showing the clustering pattern of 139 F_3 lines of 4 biparental populations of tomato based on the hierarchical clustering on principal components analysis (HCPC). Cluster D: 'pop_3_W/D1' (n = 34), Cluster H: 'pop_1_W/H1' (n = 35), Cluster T: 'pop_5_W/T1' (n = 33), and Cluster U: 'pop_7_W/U1' (n = 37).

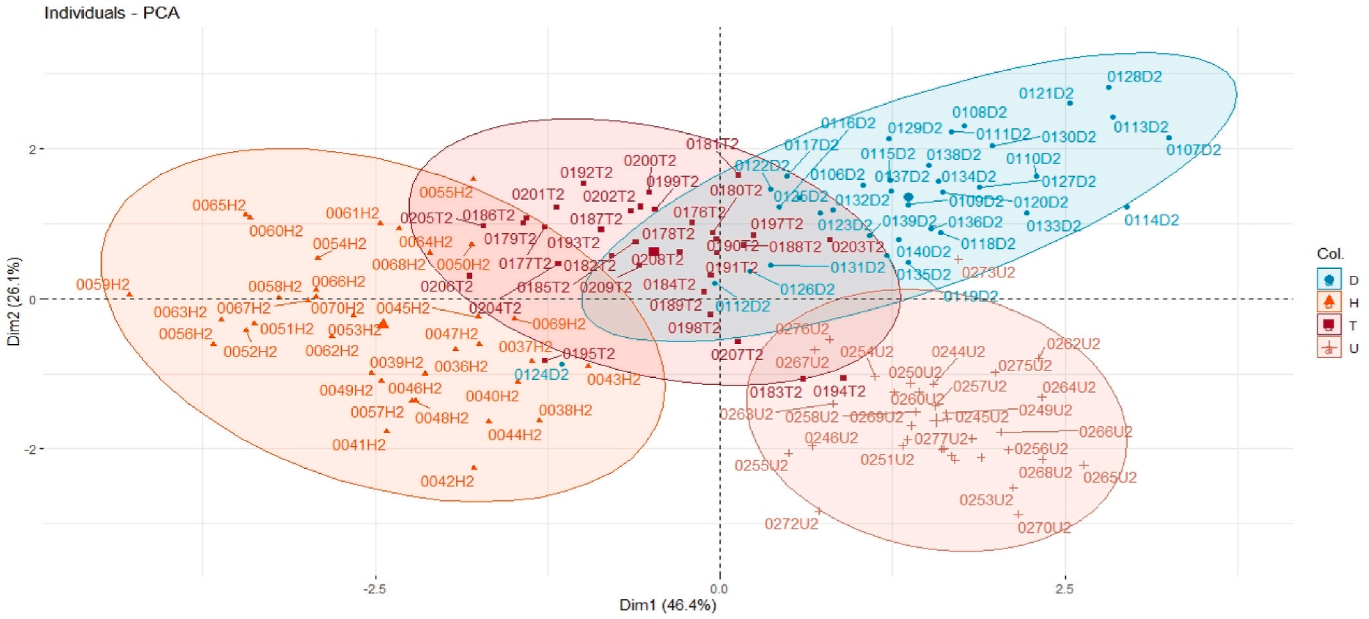

Fig. 12. Factor map showing the clustering pattern of 137 BCE_1F_2 lines of 4 biparental populations of tomato based on the hierarchical clustering on principal components analysis (HCPC). Cluster D: 'pop_4_W/D2' (n = 34), Cluster H: 'pop_2_W/H2' (n = 34), Cluster T: 'pop_6_W/T2' (n = 33), and Cluster U: 'pop_8_W/U2' (n = 36).

32.35% of the 34 original lines of cluster D and 48.48% of all T lines overlapped between clusters D and T, while 18.18% of all T lines and 8.82% of all H lines were found transgressive between clusters T 'pop_6_W/T2' and H 'pop_2_W/H2'. Very negligible number of cluster U lines was found overlapped with cluster T. Apart from all the phenological traits, NBp and NBs where cluster U performed best, cluster T took leadership for the rest yield enhancing traits including yield itself (Supplemental Tables S4b, S5b, S6b, & S7b). There was significant population structuring among the lines observed in BC_1F_2 composite population used in this study.

Some lines were found elsewhere outside their original clusters. For instance, lines '0195T2' and '0194T2' of cluster T were found in clusters H and U, respectively; line '0055H2' of cluster H drifted into cluster T, line '0124D2' of cluster D fell into cluster H and '0272U2' of cluster U was found in cluster D. BC_1F_2 lines were generated through backcross to the recurrent parents. The description of the clusters D, H, T, and U of the BC_1F_2 population is found in Fig. 12.

Using the boxplot analysis, the mean of each BC_1F_2 cluster group

showed significant differences ($p < 0.05$) for all studied yield enhancing traits (Supplemental Fig. S2). Cluster H outperformed the rest clusters in mean values for PH, TNFlPP, and TNFrPP, although it took the longest days to achieve first anthsis, 50% anthesis, and 50% fruit set. Cluster U was consistent with the least mean values for DFA, D50A, D50_FS, and DFFE, except DFFR where cluster T showed the least mean value. Cluster U 'pop_8_W/U2' also had highest mean value for NB_sec. Cluster T 'pop_6_W/T2' had the highest mean values for NN, FWPP, NLPF, and fruit yield. Clusters D and U were comparatively higher for NL, NB_prim, D1stFSp, and D100FSp.

4. Discussion

4.1. Comparative analysis of variation among the selfed and backcrossed populations

Analysis of variance showed significant variation for virtually all traits studied, as well as wider ranges for most traits in all the 8 populations including the F_3 and BC_1F_2 population types. The highest maximum range and mean values among populations for various traits gave an indication that selection for PH, TNFlPP, TNFrPP, and NLPF will favor 'pop_1_W/H1'; NL, NBp, NBs, NN, least DFA, D50A and D50FS will enhance 'pop_7_W/U1'; FWPP and fruit yield, least DFFE and DFFR will improve 'pop_5_W/T1' of selfed (F_3) populations, whereas selection for increased D1stFSp and D100FSp will favor the backcross population 'pop_8_W/U2'. Gomes et al. [20] reported the presence of variability among the BC_1F_3 populations of Santa Cruz type dwarf tomato which supported the selection of tomatoes that stood out for agronomic and quality traits. De Oliveira et al. [38] in another study reported genetic dissimilarities among 19 BCE_1F_3 populations evaluated for agronomic and fruit quality traits. In the present study, evaluating F_3 and BC_1F_2 populations simultaneously opened up opportunity for comparison, of which the selfed populations showed better variability and yield trait increased performance. Atugwu and Uguru [18] reported enormous variation and fruit size increase among early recombinants of a cross between cultivated species and *S. pimpinellifolium* accession. The high or medium to high GCV and PCV displayed for a majority of traits in one population or the other showed greater phenotypic and genotypic variability among the lines of the affected populations, hence, higher strength of the affected traits to make progress in selection. The opposite is the case for traits that expressed low GCV and PCV values. The narrow differences between PCV and GCV for all the traits in all populations showed less influence of the environment on the expression of these traits, suggesting the presence of genetic variance and hence, indicated that selection could be effective and provides desirable output for improvement [39]. Furthermore, it reflects the potential for these affected traits to respond to selection pressure on affected segregating populations, a case of evolvability of the traits [40]. The genetic variance, specifically, additive genetic variance is a mean-standardized index of the evolvability of a trait [40,41]. In other words, evolvability expresses a trait's capacity or potential to adapt to various selection pressures to which a given population whether natural or derived is subject [42]. The larger the coefficient of variance whether the genetic variance or additive genetic variance, the greater the evolvability of the trait. Heritability and genetic advance simultaneously give a more reliable gain from selection in a breeding program. High or moderate to high heritability followed by similar categories of genetic advance as percentage of mean for PH, NL, NBp, NLPF, FWPP, fruit shelf life traits, and total fruit yield observed in one population or the other indicated that the inheritance pattern in these traits could be due to additive gene actions, which also showed the effectiveness of selection to improve the traits at $F_{2\text{-}3}$ generation [43]. The phenological traits *viz.*, DFA, D50A, DFFE, D50FS, and DFFR, which expressed low values of genetic advance in a majority of the populations gave the traits a drawback for selection although the heritability for some were high or moderate to high. Low GAM indicated the involvement of non additive gene action in the inheritance of these traits, giving large scope for heterosis breeding as an improvement approach [44]. Saravanan et al. [44] reported high heritability combined with low genetic advance as a percentage of mean for days to 50 % flowering in tomato. Among the populations studied, 'pop_1_W/H1' followed by 'pop_4_W/D2' and 'pop_6_W/T2' expressed the highest genetic divergence for PH, NL, TNFlPP, TNFrPP, and FWPP based on their high magnitude of PV and GV. This revealed options or signs of future success for obtaining best fruit producing, humid tolerant tomato lines. Finzi et al. [19] identified some tomato populations with the highest divergence and the most promising for development of inbred lines with improved fruit traits among the 12 BCE_1F_2 populations of dwarf round tomatoes. Genetic dissimilarity was also reported in BC_1F_3 populations of saladette tomato [38].

4.2. Tomato selfed populations performed better than backcrossed populations

The correlation coefficient between two particular traits is approximately equal to the cosine of the angle across the space separating their vectors [45]. The positive correlation witnessed virtually for all yield*trait combinations implied that yield combined with all the other traits as indicated by the acute angles of their vectors in the biplot. Also, the long vector which represents each of the yield*trait combination indicates very strong and high relationship among the yield*trait combinations. These showed that one of the yield*traits could be enough selection criterion. A major advantage of relationship among yield*trait biplot view is its ability to identify excessive traits, select fewer traits and reduce costs in traits measurement and management without undermining experiment precision.

The trait profile experience showed excellent performance of selfed populations 'pop_5_W/T1' and 'pop_1_W/H1' for yield*TNFlPP, yield*TNFrPP, yield*PH and yield*NN, although they had high proximity with the other yield*trait combinations. Hence, these populations were considered promising, each with specific positive yield*traits associations. Relationship among yield*trait biplot view is not considered as the best to check trait profiles of the genotypes [21]. However, becoming aware of the interrelationships among the yield*trait combinations make it easier to take a well informed decision on any genetic material for breeding programs, aimed at selecting trait of interest for improvement [46–48].

The Average Tester (trait) Coordination view of GYT biplot is effective as genotypes are ranked on the account of their levels in combining yield with target traits, and simultaneously indicating the strengths and weaknesses of the genotypes [21]. The results also expressed the possibility of determining contrasting populations based on the yield*traits combinations for improving tomato breeding. In the present study, 'pop_5_W/T1', 'pop_1_W/H1', 'pop_4_W/D2', and 'pop_6_W/T2 ranked as the best in combination of some yield enhancing traits with fruit yield under cool tropical monsoon climatic conditions of Jimma, Ethiopia. This was based on the magnitude of their mean values as expressed in the mean superiority index (MSI) table, which was evidence of the yield*trait profiles of the populations. Of these, 'pop_5_W/T1' did not have any negative value for all traits, which showed that this population had the best characteristic performance across multi-traits studied; while 'pop_1_W/H1' which followed 'pop_5_W/T1' based on MSI table exhibited negative value only for D1stFSp indicating that this population had low value for this traits among the best performers. The MSI ranked the tomato populations by means of all yield*traits studied, where high values of MSI showed the best performing population (s). Other studies have used MSI ranking to identify top performers, like Yan and Frégeau-Reid [21] for Oat, Mohammadi [22] for durum wheat, and Lance et al. [49] for red spring wheat.

Using the polygon model, 'pop_5_W/T1' and 'pop_1_W/H1' also exhibited the best performance for a majority of the yield*trait combinations which showed they were the best in combining fruit yield with

the rest traits in the humid ecology. 'pop_4_W/D2' together with 'pop_8_W/U2' having taken the longest time to achieve the initial and 100% fruit spoilage proved to possess more gene for longer fruit shelf life from the wild parent 'LA2093' among the other populations. The polygon view of genotype by yield*trait (GYT) biplot has recently been used to study yield*trait combination profiles of genotypes, genotype evaluation and selection for improvement in several crop species including oat [21,50], sesame [51], durum wheat [22,48], cowpea [47, 52], peanut [53], cotton [54], and *Jatropha curcas* [55].

4.3. Clustering pattern analysis

The clustering pattern was supported by the significant differences observed among selfed and backcrossed populations for all seventeen investigated traits. This variation registered among these tomato populations/clusters may be due to differences and similarities in genetic constitution of the genotypes. This is partly in agreement with the report of Kanneh et al. [56] who noted that variation observed among tomato interspecific genotypes was due to differences in genetic and environment conditions. Interspecific hybridization can mix genetic materials from two different species, which has proven to be an effective way to increase phenotypic variability [3]. The significant structuring among the lines of F_3 and BC_1F_2 composite populations although with a bit of transgressive segregation and overlapping lines contradicts the findings of Yu et al. [57] and Bajgain et al. [58] who reported no significant population structure among the recombinant inbred lines of the nested association mapping population design. According to the authors, the lack of structure among the lines could be credited to the parents, which were diverse from each other. However, the population structure pattern observed in F_3 and BC_1F_2 composite populations in the present study followed the relatedness found among the recurrent parents in an earlier genetic characterization study [59]. The closeness in relationship among the recurrent parents could have contributed to the observed population structuring pattern among the lines.

The high percentage of overlapping lines observed between clusters D and U, as well as clusters H and T for the studied traits in F_3 populations indicated a level of close tie between the overlapping clusters. It could be suspected that the genetic distance between the parental lines that generated lines of these clusters that showed overlapping are similar. Ene et al. [59] reported the grouping of the recurrent parents of clusters D (CLN2498D) and U (UC Dan INDIA) in the same cluster while those of clusters H (CLN2417H) and T (Tima) in the same cluster showing a level of relatedness. Although each cluster was of different interspecific cross, they were half sib progenies sharing a common parent (pollen donor-wild parent 'LA2093') from the initial cross. This established a level of relationship among the progenies altogether. Similar implication could be attributed to the same lines overlapping or drifting witnessed between clusters D and T, as well as clusters T and H of BC_1F_2 populations. Rieseberg et al. [60] reported that transgressive segregation is favored by high genetic distance between the parental lines, preferably of different species.

The outliers '0210U1' and '0211U1' lines from cluster U; lines '0020H1' and '0035H1' of cluster H which found their way among lines of cluster T; and '0144T1' of cluster T which showed slight drifting into cluster H, revealed trangressive behavior of the segregating lines among selfed populations. A situation where little minority of recombinants are outliers relative to the range of parental phenotypes, and is commonly observed in plant breeding populations [61]. Whereas this phenomenon has been linked to complementary action of gene and epistasis, the biochemical, physiological, and molecular bases have not been fully understood [62]. Interspecific transgressive individuals have been argued to represent a potential source of novel genetic variation in crop species as they can potentially affect characters of adaptive significance [62–64].

Similar implication could be likened to the lines '0195T2' and '0194T2' of cluster T which found their way into clusters H and U, respectively; line '0055H2' of cluster H which drifted into cluster T, line '0124D2' of cluster D which fell into cluster H and '0272U2' of cluster U finding its way into cluster D, among BC_1F_2 populations. De Vicente and Tanksley [63] reported that interspecific transgressive lines of tomato possessed characteristics that allowed them to occupy new ecological niches where others could not or better competed in existing environments. Epistatic interactions of parental alleles and complementary action of additive alleles are regarded as being responsible for the superior or inferior attributes of transgressive segregants [65].

Among selfed (F_3) populations, desirable performance appeared for PH, TNFlPP, and TNFrPP in cluster H (pop_1_W/H1); NB_prim, NB_sec, NN, FWPP, fruit yield, early DFFE and DFFR in cluster T ('pop_5_W/T1'); and NL, D1stFSp, and D100FSp in cluster U ('pop_7_W/U1'). For BC_1F_2, desirable performance was observed in cluster U ('pop_8_W/U2') for early DFA, D50A, D50_FS, and DFFE; cluster T (pop_6_W/T2) for NN, FWPP, NLPF, and fruit yield; clusters D and U for NL, NB_prim, D1stFSp, and D100FSp. The findings suggested that improvement of these traits through single plant selection would favor the genotypes of the respective clusters and could be beneficial for future tomato breeding program, after verifying their consistencies over different environments [66]. It further advocated that high yielding and adaptable lines to high humidity could emerge from interspecific hybridization following single plant selection [67].

5. Conclusions

Genetic variability study revealed considerable dissimilarity among the lines within each population and between populations that would be useful for fruit yield improvement and adaptability to tropical humid climes. High or moderate to high heritability and high genetic advance as percentage of mean for PH, NL, NB_prim, NN, NLPF, TNFlPP, TNFrPP, FWPP, D1stFSp, D100FSp, and total fruit yield noticed in one population or the other indicated that their inheritance pattern were due to additive gene actions, which suggested selection as an effective improvement approach. Among the F_3 and BC_1F_2 populations, 'pop_1_W/H1', 'pop_4_W/D2' and 'pop_6_W/T2' which expressed the highest genetic divergence for PH, NL, TNFlPP, TNFrPP, and FWPP were most promising for development of inbred lines with improved fruit traits. Selection for increased post-harvest durability will favor 'pop_4_W/D2' and 'pop_8_W/U2'.

Application of GYT biplots models in this study presents a novel approach to tomato population improvement based on multiplex traits. 'pop_5_W/T1' and 'pop_1_W/H1' of F_3 population, ranked the best in combining some yield enhancing traits with fruit yield, indicating genetic gain and showing adaptability to the growing environment. GYT biplots could help to conquer the general challenge of crop selection on multiple traits which has been a problem in plant breeding.

Clustering pattern showed significant differences among F_3 and BC_1F_2 populations for all studied traits. Cluster T ('pop_5_W/T1') among F_3 populations, and cluster T (pop_6_W/T2) among BC_1F_2 populations showed desirable performance for fruit weight and fruit yield, and of course any improvement program for fruit yield under humid condition would favor the genotypes of these clusters. Using multivariate analysis, transgressive segregants '0210U1', '0211U1', and '0171T1' of selfed (F_3) population were observed. They are believed to represent a potential source of novel genetic variation for future tomato breeding program.

Funding

This work was supported financially by the European Union Commission under the GENEs Project – with Grant [2016-2988/001-001] (an Intra-Africa Academic Mobility Scheme). This work is part of the PhD research for C.O. Ene.

CRediT authorship contribution statement

Chikezie Onuora Ene: Conceptualization, Investigation, Methodology, Writing – original draft. **Wosene Gebreselassie Abtew:** Conceptualization, Investigation, Project administration, Supervision. **Happiness Ogba Oselebe:** Funding acquisition, Project administration, Resources, Supervision. **Uchechukwu Paschal Chukwudi:** Formal analysis, Writing – review & editing. **Emeka Chibuzor Okechukwu:** Formal analysis, Visualization, Writing – review & editing. **Friday Ugadu Ozi:** Data curation, Methodology, Visualization. **Temesgen Matiwos Menamo:** Conceptualization, Formal analysis, Methodology. **Chibueze Kelechi Ene:** Data curation, Software. **Agatha Ifeoma Atugwu:** Supervision, Visualization.

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.

Data availability

Data will be made available on request.

Acknowledgements

We are grateful to the Ethiopian Agricultural Research Institute-Melkassa Agricultural Research Center, Ethiopia; National Horticultural Research Institute-Ibadan, Nigeria and C.M. Rick Tomato Genetic Resources Center, Department of Vegetable Crops, University of California, for providing us with the planting materials.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jafr.2024.100993.

References

[1] Faostat, Global Tomato Production, 2023. http://www.fao.org/faostat/en/#data/QC. (Accessed 22 August 2023).
[2] Faostat, Global Tomato Production, 2022. http://www.fao.org/faostat/en/#data/QC. (Accessed 27 December 2022).
[3] X. Wang, L. Gao, C. Jiao, S. Stravoravdis, S.H. Prashant, S. Saha, J. Zhang, S. Mainiero, R.S. Susan, C. Catala, B.M. Gregory, A.M. Lukas, J. Vrebalov, J. G. James, S. Wu, Z. Fe, Genome of *Solanum pimpinellifolium* provides insights into structural variants during tomato breeding, Nat. Commun. 11 (2020) 5817, https://doi.org/10.1038/s41467-020-19682-0.
[4] D.W. Asfaw, Analysis of technical efficiency of smallholder tomato producers in Asaita district, Afar National Regional State, Ethiopia, PLoS One 16 (2021) e0257366, https://doi.org/10.1371/journal.pone.0257366.
[5] A.I. Atugwu, E.C. Okechukwu, V.N. Onyia, C. Chukwu, V.N. Ede, Studies on adaptability of advanced generations of wild and cultivated tomato crosses in a humid environment, Researchjournali's Journal of Agriculture 6 (2019) 1–13.
[6] M.I. Uguru, A.I. Atugwu, Comparative study on the somatic chromosome number, growth, yield and disease incidence of cultivated tomatoes and their wild relative, Agro-Science 1 (2000) 52–58, https://doi.org/10.4314/as.v1i2.1455.
[7] H.L. Xu, Effect of ambient humidity on physiological activities and fruit yield and quality of greenhouse tomato, Acta Hortic. 761 (2007) 85–92, https://doi.org/10.17660/ActaHortic.2007.761.9.
[8] M. Alam, N. Sultana, S. Ahmad, M. Hossain, A. Islam, Performance of heat tolerant tomato hybrid lines under hot, humid conditions, Bangladesh J. Agric. Res. 35 (2010) 367–373, https://doi.org/10.3329/bjar.v35i3.6442.
[9] J.C. Miller, S.D. Tanksley, RFLP analysis of philogenetic relationships and genetic variation in the genus *Lycopersicon*, Theor. Appl. Genet. 80 (1990) 437–448, https://doi.org/10.1007/BF00226743.
[10] M.R. Foolad, Genome mapping and molecular breeding of tomato, Int. J. Plant Genom. (2007) 1–52, https://doi.org/10.1155/2007/64358.
[11] G.R. Rodriguez, G.R. Pratta, R. Zorzoli, L.A. Picardi, Recombinant lines obtained from an interspecific cross between *Lycopersicon* species selected by fruit weight and fruit shelf life, J. Am. Soc. Hortic. 131 (2006) 651–656.
[12] G.R. Rodrigues, G.R. Pratta, D.R. Liberatti, R. Zorzoli, L.A. Picardi, Inheritance of shelf life and other quality traits of tomato fruit estimated from F_1's, F_2's and backcross generations derived from standard cultivars, nor homozygote and wild cherry tomato, Euphytica 176 (2010) 137–147, https://doi.org/10.1007/s10681-010-0241-9.
[13] D. Schwarz, A.J. Thompson, H.P. Kläring, Guidelines to use tomato in experiments with a controlled environment, Front. Plant Sci. 5 (2014) 1–16, https://doi.org/10.3389/fpls.2014.00625.
[14] S. Grandillo, P. Termolino, E. Van der Knaap, Molecular mapping of complex traits in tomato, in: B.E. Liedl, J.A. Labate, J.R. Stommel, A. Slade, C. Kole (Eds.), Genetics, Genomics, and Breeding of Tomato, CRC Press; Science Pub, Boca Raton, FL, 2013, pp. 150–227.
[15] I. Celik, N. Gurbuz, A.T. Uncu, A. Frary, S. Doganlar, Genome-wide SNP discovery and QTL mapping for fruit quality traits in inbred backcross lines (IBLs) of *Solanum pimpinellifolium* L. using genotyping by sequencing, BMC Genom. 18 (2017) 1, https://doi.org/10.1186/s12864-016-3406-7.
[16] L. Piosik, M. Ruta-Piosik, M. Zenkteler, E. Zenkteler, Development of interspecific hybrids between *Solanum lycopersicum* L. and *S. siymbriifolium* Lam. via embryo calli, Euphytica 215 (2019) 31, https://doi.org/10.1007/s10681-019-2358-9.
[17] A.W. Ebert, Y.Y. Chou, The tomato collection maintained by AVRDC – the World Vegetable Center: composition, germplasm dissemination and use in breeding, Acta Hortic. 1101 (2015) 169–176, https://doi.org/10.17660/ActaHortic.2015.1101.26.
[18] A.I. Atugwu, M.I. Uguru, Tracking fruit size increase in recombinants obtained from an interspecific cross between cultivated tomato (*Solanum esculentum* L.) and wild tomato relative (*S. pimpinellifolium* L.), J. Plant Breed. 4 (2012) 62–71, https://doi.org/10.5897/JPBCS.9000131.
[19] R.R. Finzi, G.M. Maciel, H.G. Peres, M.F. Silva, J.V.M. Peixoto, D.A. Gomes, Agronomic potential of BC_1F_2 dwarf round tomato populations, Cienc. E Agrotecnol 44 (2020) e028819, https://doi.org/10.1590/1413-7054202044028819.
[20] D.A. Gomes, G.M. Maciel, A.C. S Siquieroli, C.S. De Oliveira, R.R. Finzi, D. J. Marques, Selection of BC_1F_3 populations of Santa Cruz type dwarf tomato plant by computational intelligence techniques, Bragantia Campinas 80 (2021) e4821, https://doi.org/10.1590/1678-4499.20210046.
[21] W. Yan, J. Frégeau-Reid, Genotype by yield*trait (GYT) biplot: a novel approach for genotype selection based on multiple traits, Sci. Rep. 8 (2018) 8242, https://doi.org/10.1038/s41598-018-26688-8.
[22] R. Mohammadi, Genotype by yield*trait biplot for genotype evaluation and trait profiles in durum wheat, Cereal Res. Commun. 47 (2019) 541–551, https://doi.org/10.1556/0806.47.2019.32.
[23] N. Abdela, B.F. Deressa, A. Hassan, E. Teshome, Prevalence of indigestible foreign bodies in the rumen and reticulum of sheep slaughtered at Jimma municipal abattoir, Southwestern Ethiopia, Global J. Med. Res. 17 (2017) 27–34.
[24] S. Alemineh, S. Senadheera, Y. Birru, Engineering characterization of subgrade soils of Jimma town, Ethiopia, for roadway design, Geosci. 10 (2020) 94, https://doi.org/10.3390/geosciences10030094.
[25] M. Ozores-Hampton, Hand Pollination of Tomato for Breeding and Seed Production, UF/IFAS Southwest Florida Research and Education Center, Immokalee, FL, 2014, 34142. HS1248, EDIS, http://edis.ifas.ufl.edu.
[26] M.K. Osei, R. Akromah, S.L. Shilh, S.K. Green, Evaluation of some tomato germplasm for resistance to tomato yellow leaf curls virus disease (TYLCV) in Ghana, Aspect Appl. Biol. 96 (2010) 315–323.
[27] F.F. Dinssa, O. Mbwambo, J. Macharia, A. Aloyce, O. Shilla, Protocols for fruit, leafy and bulb-type vegetable cultivar testing, in: AVRDC – the World Vegetable Center. Shanhua, Taiwan. Publication 15-793, 2015, p. 19.
[28] I.K. Arah, H. Amaglo, E.K. Kumah, H. Ofori, Preharvest and postharvest factors affecting the quality and shelf life of harvested tomatoes, Int. J. Agron. (2015) 1–6, https://doi.org/10.1155/2015/478041.
[29] K.A. Gomez, A.A. Gomez, Statistical Procedures for Agricultural Research, second ed., An international rice research institute book. A Wiley-interscience Publication. John Wiley & Sons, New York, 1983, p. 657. Chichester. Brisbane. Toronto. Singapore.
[30] R Core Team, A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, 2017. http://www.R-project.org/.
[31] P. Raj, P. Rumit, P. Dinesh, Variability: Genetic Variability Analysis for Plant Breeding Research, R Package Version, 2020. http://cran.r-project.org/package=variability.
[32] G.W. Burton, E.H. De Vane, Estimating heritability in tall fescue (*Festuca arundinacea* L) from replicated clonal material, Agron. J. 45 (1953) 478–481, https://doi.org/10.2134/agronj1953.00021962004500100005x.
[33] R.K. Singh, B.D. Choudhary, Biometrical Methods in Quantitative Genetic Analysis, Kalyani, New Delhi, India, 1985.
[34] M.M.H. Khan, M.Y. Rafii, I.S. Ramlee, M. Jusoh, A.I. Mamun, Genetic variability, heritability, and clustering pattern exploration of Bambara groundnut (*Vigna subterranean* [L] Verdc.) accessions for the perfection of yield and yield-related traits, BioMed Res. Int. (2020) 2195797, https://doi.org/10.1155/2020/2195797.
[35] W. Yan, GGE biplot-a windows application for graphical analysis of multi-environmental trial data and other types of two-way data, Agron. J. 93 (2001) 1111–1118, https://doi.org/10.2134/agronj2001.9351111x.
[36] W. Yan, N.A. Tinker, A biplot approach for investigating QTL-by-environment patterns, Mol 15 (2005) 31–43, https://doi.org/10.1007/s11032-004-1706-0.
[37] A. Kassambara, F. Mundt, Factoextra: extract and visualize the results of multivariate data analyses, in: R Package Version 1, vol. 7, 2020 ed2020, 0.
[38] C.S. De Oliveira, G.M. Maciel, A.C.S. Siquieroli, D.A. Gomes, N.M. Diniz, J.M. Q. Luz, R.Y. Yada, Artificial neural networks and genetic dissimilarity among saladette type dwarf tomato plant populations, Food Chem. 3 (2021) 100056, https://doi.org/10.1016/j.fochms.2021.100056.

[39] M.Y. Saleem, Q. Iqbal, M. Asghar, Genetic variability, heritability, character association and path analysis in F_1 hybrids of tomato, Pak. J. Agri. Sci. 50 (2013) 649–653. http://www.pakjas.com.pk.
[40] W.G. Hill, Understanding and using quantitative genetic variation, Philos. Trans. R. Soc. 365 (2010) 73–85, https://doi.org/10.1098/rstb.2009.0203.
[41] D. Houle, Comparing evolvability and variability of quantitative traits, Genetics 130 (1992) 195–204.
[42] B.Y. Cheung, Genetic coefficient of variance, in: V. Zeigler-Hill, T.K. Shackelford (Eds.), Encyclopedia of Personality and Individual Differences, Springer, Cham, 2020, 10/1007/978-3-319-24612-3_1474.
[43] M. Henareh, Genetic variation in superior tomato genotypes collected from Northwest of Iran, Int'l. J. Scientific. Res. Environ. Sci. 3 (2015) 219–225.
[44] K.R. Saravanan, V. Vishnupriya, M. Prakash, R. Anandan, Variability, heritability and genetic advance in tomato genotypes, Indian J. Agric. Sci. 53 (2019) 92–95, https://doi.org/10.18805/IJARe.A-5030.
[45] W. Yan, M.S. Kang, GGE Biplot Analysis: A Graphical Tool for Breeders, Geneticists, and Agronomist, CRC press, Boca Raton, FL, 2003, p. 288, https://doi.org/10.1201/9781420040371.
[46] M.S.D. Mendonça, P.M. Beber, F.S.S.D. Nascimento, V.B.D. Santos, J.T. Marinho, Importance and correlations of characters for cowpea diversity in traditional varieties, Rev. Cienc. 49 (2018) 267–274, https://doi.org/10.5935/1806-6690.20180030.
[47] T.R.A. De Oliveira, A. Gravina Gde, M. Rocha Mde, A. Neto Fde, D.P. da Cruz, G.H. F. de Oliveira, S.S. de Sant'Anna Cqda, C. Jaeggi Mepda, R.S. Rocha, GYT biplot analysis: A new approach for cowpea line selection, J. Exp. Agric. Int. 41 (2019) 1–9, https://doi.org/10.9734/JEAI/2019/v41i530408.
[48] E. Kendal, Comparing durum wheat genotypes by genotype × yield × trait and genotype × trait biplot method, Chil. J. Agric. Res. 79 (2019) 512–522, https://doi.org/10.4067/S0718-58392019000400512.
[49] F.M. Lance, D.G. Karl, D. Yabwalo, Use of genotype by yield*trait (GYT) analysis to select hard red spring wheat with elevated performance for agronomic and disease resistance traits, Crop Breed, Genet Genom 2 (2020) e200009, https://doi.org/10.20900/cbgg20200009.
[50] W. Yan, J. Fregeau-Reid, N. Mountain, J. Kobler, Genotype and management evaluation based on genotype by Yield*Trait (GYT) analysis, crop breed, Genet Genom 1 (2019) e190002, https://doi.org/10.20900/cbgg20190002.
[51] S. Boureima, A. Yaou, Genotype by yield*trait combination biplot approach to evaluate sesame genotypes on multiple traits basis, Turkish J, Field Crop 24 (2019) 237–244, https://doi.org/10.17557/tjfc.655165.
[52] K.C. Araújo, T.R.A. de Oliveira, A. Gravina Gde, M. Rocha Mde, A. Neto Fde, G.H. F. de Oliveira, D.P. da Cruz, S.S. de Sant'Anna Cqda, L.M. Gravina, R.S. Rocha, I. M. Pereira, Selection of high-performance black-eyed cowpea (*Vigna unguiculata* L.) through GYT biplot analysis: a new methodology based on multiple characteristics, Aust. J. Crop. Sci. 15 (2021) 464–469, https://doi.org/10.21475/ajcs.21.15.03.p2526.
[53] M.W.S. Mahmoud, E.M.A. Hussein, K.M. AboEl-kassem, H.E.A. Ibrahim, Graphical presentation of some peanut genotypes by comparing two patterns of biplot analysis, JPP 11 (2020) 697–705, https://doi.org/10.21608/jpp.2020.112895.
[54] M.A. Peixoto, J.S.P.C. Evangelista, I.F. Coelho, L.P. Carvalho, F.J.C. Farias, P. E. Teodoro, L. Bhering, Genotype selection based on multiple traits in cotton crops: The application of genotype by yield*trait biplot, Acta Sci. Agron. 44 (2022) e54136, https://doi.org/10.4025/actasciagron.v44i1.54136.
[55] R.D. Purwati, T.D.A. Anggraeni, M. Machfud, Genotype by yield*trait biplot analysis to evaluate *Jatropha curcas* genotypes based on multiple traits, the 2nd International Conference on Sustainable Plantation, in: IOP Conf. Series: Earth and Environmental Science, vol. 974, 2022 012013, https://doi.org/10.1088/1755-1315/974/1/012013, 2022.
[56] S.M. Kanneh, M.K. Osei, R. Akromah, J. Gyau, Generation mean analysis of yield and yield components of early generations of interspecific crosses of tomato (*Solanum lycopersicum* L.), Int. J. Plant Breed. Genet. 10 (2016) 98–103.
[57] J. Yu, J.B. Holland, M.D. McMullen, E.S. Buckler, Genetic design and statistical power of nested association mapping in maize, Genetics 178 (2008) 539–551, https://doi.org/10.1534/genetics.107.074245.
[58] P. Bajgain, M.N. Rouse, T.J. Tsilo, G.K. Macharia, S. Bhavani, Y. Jin, et al., Nested association mapping of stem rust resistance in wheat using genotyping by sequencing, PLoS One 11 (2016) e0155760, https://doi.org/10.1371/journal.pone.0155760.
[59] C.O. Ene, W.G. Abtew, H.O. Oselebe, F.U. Ozi, U.N. Ikeogu, Genetic characterization and quantitative trait relationship using multivariate techniques reveal diversity among tomato germplasm, Food Sci. Nutr. 10 (2022) 2426–2442, https://doi.org/10.1002/fsn3.2850.
[60] L.H. Rieseberg, M.A. Archer, R.K. Wayne, Transgressive segregation, adaptation and speciation, Heredity 83 (1999) 363–372.
[61] E.A. Boyle, Y.I. Li, J.K. Pritchard, An expanded view of complex traits: from polygenic to omnigenic, Cell 169 (2017) 1177–1186, https://doi.org/10.1016/j.cell.2017.05.038.
[62] I.C.M. Pabuayon, A. Kitazumi, K.R. Cushman, R.K. Singh, G.B. Gregorio, B. Dhatt, M. Zabet-Moghaddam, H. Walia, B.G. de los Reyes, Novel and transgressive salinity tolerance in recombinant inbred lines of rice created by physiological coupling-uncoupling and network rewiring effects, Front. Plant Sci. 12 (2021) 615277, https://doi.org/10.3389/fpls.2021.615277.
[63] M.C. De Vicente, S.D. Tanksley, QTL analysis of transgressive segregation in an interspecific tomato cross, Genetics 134 (1993) 585–596, https://doi.org/10.1093/genetics/134.2.585.
[64] M. Singh, S. Rani, N. Malhotra, G. Katna, A. Sarker, Transgressive segregations for agronomic improvement using interspecific crosses between C. arietinum L. x C. reticulatum Ladiz and C. arietinum L. x C. echinospermum Davis species, PLoS One 13 (2018) e0203082, https://doi.org/10.1371/journal.pone.0203082.
[65] D.R. Dittrich-Reed, B.M. Fitzpatrick, Transgressive hybrids as hopeful monsters, Evol. Biol. 40 (2013) 310–315, https://doi.org/10.1007/s11692-012-9209-0.
[66] M.J. Gonzalo, Y.C. Li, K.Y. Chen, D. Gil, T. Montoro, I. Nájera, C. Baixauli, A. Granell, A.J. Monforte, Genetic control of reproductive traits in tomatoes under high temperature, Front. Plant Sci. 11 (2020) 326, https://doi.org/10.3389/fpls.2020.00326.
[67] P.M. Gaur, N. Mallikarjuna, T. Knights, T.S. Beebe, D. Debouck, Gene introgression in grain legumes, in: S. Gupta, M. Ali, B.B. Singh (Eds.), Grain Legumes: Genetic Improvement, Management and Trade, Indian Society of Pulses Research and Development, Indian Institute of Pulses Research, Kanpur, India, 2008, pp. 1–17.