Saudi Journal of Biological Sciences 28 (2021) 4833–4844

Contents lists available at ScienceDirect

Saudi Journal of Biological Sciences

journal homepage: www.sciencedirect.com

Original article

Genetic diversity of Indian garlic core germplasm using agro-biochemical traits and SRAP markers

Ashwini Prashant Benke [a,*], Ram Krishna [a], Vijay Mahajan [a], Waquar Akhter Ansari [b], Amar Jeet Gupta [a], Anil Khar [c], Poonam Shelke [a], A. Thangasamy [a], T. P. Ahammed Shabeer [d], Major Singh [a], Kiran P. Bhagat [e], Dalasanuru Chandregowda Manjunathagowda [a,*]

[a] ICAR-Directorate of Onion and Garlic Research, Rajgurunagar–410505, Pune, Maharashtra, India
[b] Department of Botany, Savitribai Phule Pune University, Pune 41100, Maharashtra, India
[c] Department of Vegetable Science, ICAR-Indian Agriculture Research Institute, New Delhi 110012, India
[d] ICAR-National Research Centre for Grapes, Pune 412307, Maharashtra, India
[e] ICAR-Directorate of Floriculture, Pune 411005, Maharashtra, India

ARTICLE INFO

Article history:
Received 28 March 2021
Revised 2 May 2021
Accepted 3 May 2021
Available online 13 May 2021

Keywords:
Garlic
SRAP marker
Diversity analysis
Allicin

ABSTRACT

The characterization of garlic germplasm improves its utility, despite the fact that garlic hasn't been used much in the past. Garlic has an untapped genetic pool of immense economic and medicinal value in India. Hence, using heuristic core collection approach, a core set of 46 accessions were selected from 625 Indian garlic accessions based on 13 quantitative and five qualitative traits. The statistical measures (CV per cent, CR per cent, VR per cent) were used to sort the core set using Shannon-Wiener diversity index and the Nei diversity index. In addition, the variation within the core set was tested for 18 agro-morphological and six biochemical characteristics (allicin, phenol content, pyruvic acid, protein, allyl methyl thiosulfinate (AMTHS), and methyl allyl thiosulfinate (MATHS)). Further study of the core set's molecular diversity was performed using sequence related amplified polymorphism (SRAP) markers, which revealed a wide range of diversity among the core set's accessions, with an average polymorphism efficiency (PE) of 80.59 percent, polymorphism information content (PIC) of 0.29, effective multiplex ratio (EMR) of 3.51, and marker index (MI) of 0.99. The findings of this study will be useful in identifying high-yielding, elite garlic germplasm lines with the trait of interest. Since this core set is indicative of total germplasm, these selected breeding lines will be used for genetic improvement of garlic in the future.

1. Introduction

The Alliaceae family, which includes essential *Allium* crops such as onion, garlic, leek, and chives, is the second most important monocot family after Poaceae. Garlic (*Allium sativum* L.) is the most essential crop of the genus *Allium* after onion (*Allium cepa* L.) (Benke et al., 2020a; Khandagale et al., 2020) and thought to have arisen in Central Asia (Vavilov, 1951) India has with over 5000 years cultivation history (Benke et al., 2020a). Garlic average yield has risen from 3.48 to 5.27 MT per ha over the last 40 years in India (FAOSTAT, 2018). Garlic, on the other hand, has a lower efficiency in genetic improvement than onion, owing to sexual sterility, as a consequence, do not produce seeds, hence cloves are used to asexually propagation (Manjunathagowda et al., 2017). In spite of fertile garlic accessions have been found near the Tien Shin Mountains, the vast majority of the world's garlic genetic resources are non-flowering (Etoh and Simon, 2002). Garlic clones that do not produce flowering bolts are considered to be soft-necked garlic varieties, however, hard-neck garlic types do produce flowering bolts on rare occasions, the flowers rose in an umbel but did not set the seeds due to underdeveloped gametophytes, which may cause male and female sterility (Benke et al., 2020b). As a result, garlic is only propagated by cloves and has restricted access to conventional breeding techniques (Benke et al., 2020a). As a consequence of these features, there is a narrow genetic base with

* Corresponding authors.
E-mail addresses: ashwini.ashwini.chavan@gmail.com (A.P. Benke), Vijay.Mahajan@icar.gov.in (V. Mahajan), amarjeet.gupta@icar.gov.in (A.J. Gupta), poonam.shelke@icar.gov.in (P. Shelke), a.thangasamy@icar.gov.in (A. Thangasamy), dcmgowda@gmail.com (D.C. Manjunathagowda).
Peer review under responsibility of King Saud University.

Production and hosting by Elsevier

https://doi.org/10.1016/j.sjbs.2021.05.013
1319-562X/

little variance in various traits of garlic. Although garlic has a considerable amount of genetic variation in terms of morphological and biochemical characteristics (Zhao et al., 2011; Cunha et al., 2012).

Garlic is abundant in bioactive compounds, including organosulfur compounds such as allicin, diallyl thiosulfonate (allicin), diallyl sulfide (DAS), diallyl disulfide (DADS), diallyl trisulfide (DATS), E/Z-ajoene, S-allyl-cysteine (SAC), and S-allyl-cysteine sulfoxide (alliin) are the key sulfur compounds (Khar et al., 2011). As a result, all of these factors offer a range of health benefits and used to treat stomach issues, respiratory problems, parasitic infestations, diabetes, high cholesterol, leprosy, and a number of other ailments (Khandagale et al., 2020; Benke et al., 2019) and its intake in the diet can help to prevent cancer, heart disease, and neurodegenerative diseases (Chand et al., 2015). Upon importance of organosulfur compounds, identifying genotypes with high concentrations of bioactive compounds in the available ecotypes of Indian garlic gene bank would be essential for human health (Khar et al., 2011) These organosulfur compounds (pungency and flavours) in genotypes (Abedi et al., 2013) vary with genotype and environment, thus it is necessary to study the genetic variation among garlic accessions. It has been studied using multilocus and dominant molecular markers such as inter simple sequence repeat (ISSR) (Chand et al., 2015) random amplified polymorphic DNA (RAPD) (Abdoli et al., 2009) and amplified fragment length polymorphisms (AFLP) (Morales et al., 2013). At the taxonomic level, all markers display low polymorphism, including single sequence repeats (SSR) (Cunha et al., 2012). Recently, used sequence related amplified polymorphism (SRAP), found to be as accurate and variable as AFLP though requiring less technological expertise (Chen et al., 2013). SRAP markers amplify the coding region of DNA and target the open reading frame of the genome (Robarts and Wolfe ., 2014).

The Indian garlic field gene bank comprises 625 ecotypes, depicting a wide variety of agro-morphological traits. It takes a lot of time and resources to phenotype and genotype this entire collection at the molecular and biochemical level for long-term storage in a gene bank (Khandagale et al., 2020). Working with a smaller germplasm collection will allow assessing even complex traits and making efficient use of plant germplasm far simpler (Chen et al., 2014; Kumar et al., 2019). The core selection has been a reasonable option for preserving a set of germplasm that represents the genetic diversity of the base collection, and it can now be freed from its restricted usage for evaluation. This strategy was implemented to minimize the expense of maintaining a large number of germplasm, also preserving overall genetic diversity in the germplasm chosen. This collection decreases the maintenance burden (*in vivo* or *in vitro*) in the event of natural disasters or pandemics. To date, several sampling strategies for core set creation have been successful, including random sampling, deviation sampling, and preferred sampling, all of which maintain the base collection's maximum genetic variation structure (Hu et al., 2000). The unweighted pair-group average method (Hua et al., 2017), Ward's method (Ward, 1963) full linkage method of hierarchical clustering (Hu et al., 2000) and heuristic core set selection (HCC) (Gowda et al., 2012) approaches were used to determine group boundaries while constructing the core subset. However, the South Korean Rural Development Administration (RDA) recently created software called "Power core" (Kim et al., 2007) (http://genebank.rda.go.kr/powercore) that can be used to sort the core community. Considering the importance of core set extraction, the concept has been already applied worldwide in rice (Chung et al., 2009), wheat (Dutta et al.,2015), foxtail millet (Gowda et al, 2012), bean (Vaijayanthi et al., 2015), pigeon pea (Upadhyaya et al., 2006), soybean (Oliveira et al., 2010), cashew (Mohana and Nayak, 2018), groundnut (Upadhyaya et al., 2003), brinjal (Gangopadhyay et al., 2010). In this context, we established the garlic core set from 625 germplasm collection of the Indian garlic gene bank at the national active germplasm site, and further we characterize the core set collection for genetic, agronomical and biochemical traits .

2. Materials and methods

2.1. Garlic germplasm used to construct the core set

2.1.1. Planting material

A total of 625 garlic accessions (Table S1), including landraces, cultivars, elite lines, and improved varieties, were evaluated for 13 quantitative and 5 qualitative characteristics, and the results were used to establish a core set. The field experiment was laid out in a Latin square design (LSD) (25 row and 25 columns), with a plot size 1 × 1.2 m for each accession. In each plot, total seven lines plotted on one meter length separated by 15 cm and ten cloves planted per line at 10 cm distance, such a way total 70 cloves of each genotypes planted in each replication. Each accession was planted in two replicates. After 70 days of planting, field observations were made for growth traits, and bulb-related traits were noted after 30 days of harvesting (DTH). The experimental trial was conducted in the winter cropping seasons of 2014–15 and 2015–16 at the Directorate of Onion and Garlic Research (DOGR), Rajgurunagar, Pune, India (27°19′00.2″N 82°25′00.1″E) working under the ages of the Indian Council of Agriculture Research (ICAR).

2.1.2. Development of core set garlic germplasm

For core set sorting, average mean data of 625 garlic accessions for recorded qualitative and quantitative traits examined through "Power core" by selecting random selection and heuristic core set approach separately. Software showed graphical representation for each trait regarding grouping of entire accessions and respective selected core entries. At end of the program, core set selected entries along with different statistical indices for qualitative and quantitative traits were deployed for "Power core" software for core formation (Kim et al., 2007).

2.1.3. Experimental details of core set characterization

The garlic core set accessions (Table S2) were planted in two replications in a randomized block design (RBD) during the *rabi* seasons of 2016–17 and 2017–18 respectively at the ICAR-DOGR. These genotypes were evaluated for 15 quantitative and three qualitative agro-morphological traits, as well as six biochemical traits (phenol, protein, pyruvic acid, allicin, AMTHS, and ATPTHS) and genetic diversity was assay using SRAP molecular markers.

2.1.4. Characterization of cores set for agronomic traits

The vegetative growth traits such as plant height (cm), pseudo stem length (cm), pseudo stem diameter (mm), number of leaves per plant, fourth leaf length (cm), and fourth leaf width (mm) were measured at 70 days after planting of cloves in experimental plot. The plant architecture and bolting behaviour (floral stalk) were noted during plant growth period as traits prescribed by the distinctness, uniformity and stability (DUS) test guidelines of National Bureau of Plant Genetic Resources (NBPGR) in New Delhi, India (www.plantauthority.in). The days to maturity of accessions recorded based on lodging, bulb traits were recorded 30 days after harvest, bulbs yield per plot considered as total yield, marketable yield (q/ha) base on the bulbs having a polar diameter (greater than 50 mm), bulb weight (g) in average of five bulbs, cloves per bulb, weight of 50 cloves (g), polar diameter (mm) and equatorial diameter (mm) of the bulb, bulb shape, bulb skin colour, and clove skin colour were noted over five representative samples per accession in both replicates of the studied years.

2.2. Analysis of biochemical compounds in the core set

In each accession, cloves of five separate bulbs were used in each replication, three biological replicates were used to test biochemical traits.

2.2.1. Estimation of phenol and protein

The total phenol content was determined (Folin and Denis, 1915), with minor modifications, in which extracts were diluted five times with distilled water after extracting five grams of bulb tissue with 70% hot ethanol. One ml of diluted extract was combined with one millilitre (ml) of 1 N phenol reagent (Wako Pure Chemical Industries, Ltd., Osaka, Japan). After three minutes, one ml of aqueous sodium carbonate solution containing ten per cent sodium carbonate was added, and the mixture was incubated at room temperature for 60 min. A U-2001 spectrophotometer was used to estimate the absorbance at 530 nm (Hitachi High-Technologies Corporation, Tokyo, Japan). The catechol calibration curve was used to measure the phenol material (Singleton and Rossi, 1965; He, 2011), which was then measured using the BSA standard curve and expressed in mg/ml.

2.2.2. Quantification of pyruvate content

The amounts of pyruvate in garlic cloves were determined using a spectrophotometer and a colorimetric technique based on the reaction of 2,4-dinitrophenylhydrazone (DNPH) (0.63 mM DNPH reagent in 0.5 mol L^{-1} HCl) with pyruvic acid. Standard solutions were made with sodium pyruvate at concentrations ranging from 0.01 to 0.2 mol ml^{-1} (Schwimmer and Weston, 1961). Shortly, 10 ul standard or sample added to 90 ul to DNPH and incubated at 25°c for 30 min. Later, 50 ul of KOH (5 mol L^{-1}) was added and incubated at 37 °c for 30 min. Eventually absorbance was recorded at 540 nm and the concentration expressed as umol of sodium pyruvate per 100 g of fresh weight as per equation y = 0.161x + 0.006 (R^2 = 0.999) produced by sodium pyruvate concentrations.

2.2.3. Determination of Allicin, AMTHS, and ATPTHS contents

The organosulphur compounds such as Allicin, AMTHS, and ATPTHS were estimated using LC-MS/MS system (API 5500 Q-trap LC-MS/MS, AB Sciex, Canada) as described (Khar et al., 2011), at the NRL, NRC on Grapes, Pune, India. Briefly, for allicin standard preparation readymade allicin tablets diluted with water to create standards for a linearity curve covering approximately 5-80 ug/ml for high performance liquid chromatoghraph (HPLC) analysis. The standards maintained at cold temperature until used. The extract from 0.7 to 1.0 g garlic samples made prepared for HPLC analysis. The analysis were made through HPLC with reversed-phase column containing an LC 300 psi gradient pump equipped with a degasifier, microwave injector, and Shimadzu (SPD-M10A) diode array detector with a wavelength of 240 nm and LC 18 reversed phase column operating at ambient temperature. The mobile phase comprises methanol and water and flow rate of 1.0 ml/min. The sample running time was 20 min.

2.3. Genetic diversity assessment using SRAP markers and analysis

SRAP molecular markers were used to assess the genetic diversity of the core set garlic accessions. A total of thirty primer combinations were created and used for the genetic diversity analysis. SRAP markers were used in the study (Table S3) (Li et al., 2007). The PCR in 25 µl reaction volume mixture has 1.5 µl template DNA (80 ng), 2.5 µl 10 × PCR buffer, 1.0 µl of dNTPs (25 mM), SRAP primer (10 pM) 1.0 µl each, *Taq* polymerase 1.5 µl and 16.5 µl sterile distilled water, the reaction was set in a VeritiTM 96 well thermal cycler (Applied Biosystems, USA). The PCR program was scheduled to start with a four minutes denaturation at 94 °C, followed by 38 cycles of one minute denaturation at 94 °C, one minute annealing at 55 °C, and one minute extension at 72 °C, with a ten minutes final extension at 72 °C. The PCR amplification with each SRAP primer combination was repeated twice to ensure reproducibility and the results were consistent. To examine the amplified SRAP amplicons, electrophoresis in 2.5 percent agarose gel was performed. The number of amplified amplicons was determined using a gel BIO-PRINT-ST4 (VilberLourmat, France) documentation system.

The SRAP amplified amplicons on gels were scored based on their absence (0) or presence (1) and the similarity matrix data was analyzed by using the Jaccard index with the help of NTSYSpc Version 2.11 × software. The polymorphism information content (PIC) was determined using the equation PICi = 2fi(1–fi) (Robarts and Wolfe, 2014) for every locus, where PICi determine the polymorphic information content of the locus I, and fi, the amplified fragments frequency, and 1-fi is the frequency of non-amplified fragments. The similarity matrix data were fed into the unweighted pair group method for the arithmetic average (UPGMA) cluster analysis tool to produce a dendrogram.

2.4. Data analysis

The mean phenotypic data of the core set were analysed using Statistical Analysis Software package v 9.3 (SAS 9.3) for univariate analysis (Kim et al, 2007). Significant variations in accessions by year were checked at a p-value of 0.05 per cent using the program proc GLM, and mean values were compared using a critical difference (CD) value of 5%. The program "Power core" was used to create the core group from overall pooled mean data. As statistical indicators, mean difference (MD), variance difference (VD), coincidence rate (CR), and variable rate (VR) in percentage were mentioned, where MD and VD describe the percentage difference between the base collection and the core subset, while core collection properties concerning the entire collection were represented (Chung et al., 2009).

The Shannon-Wiener Diversity Index and the Nei Diversity Index for qualitative traits were identified through the "Power core" program (Dutta et al., 2015). The software SAS v 9.3 of module Proc FASTA CLUS, was used to perform clustering analysis on the core set mean phenotypic and biochemical data. The dissimilarity between accessions was assessed using the "Ward method" (Ward, 1963) of clustering, and other multivariate analysis was done by the software "JMP pro 10".

3. Results

3.1. Analysis of variance

The analysis of variance of whole set revealed that, there was a non-significant association of genotypes × years at p less than 0.01, and p less than 0.05 for all traits. The coefficient of variation of various parameters had a low range value (9.82–30.39 per cent), indicating that the trial was carried out with sufficient precision (Fig. 1).

3.2. Extraction and validation of core set

The average data of 18 morphological traits from 625 garlic accessions was used for core set development (Table S1). The heuristic core collection (HCC) approach sorted 46 accessions, and the indiscriminate selection method was used to isolate 46 accessions as a core collection. CR% obtained for both core subsets was recorded to be greater than 80% (almost 98%). Core sets possessed VD% and VR% values as 56.12% and 150.18% for HCC

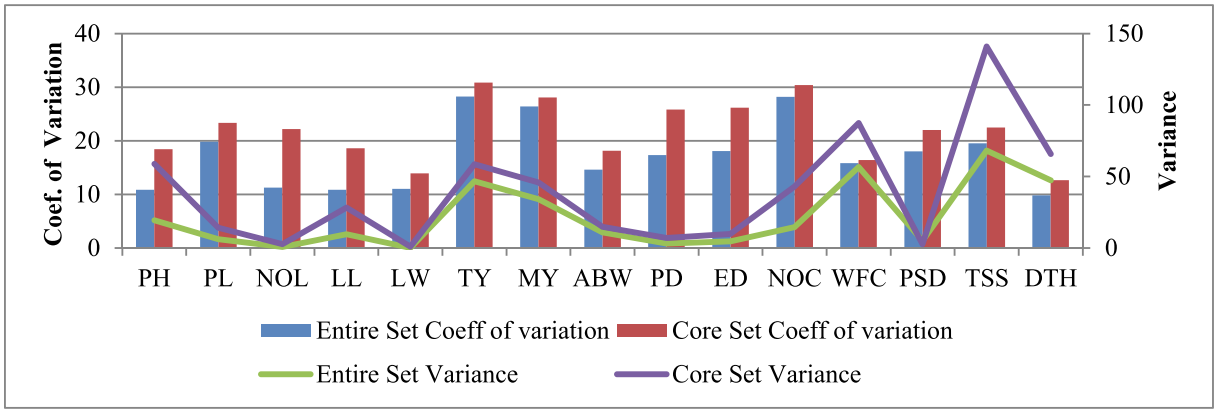

Fig. 1. Comparison between entire set and core set with respect to coefficient of variation and variance.

and 50.57% and 140.26% for the random selection method, respectively. In the present study, higher VD% compared to the random method with more precision and efficiency recorded (Table 1). Additionally, coverage percent for all the evaluated characters was 100% in the HCC approach

3.3. Comparison of entire collection and core set by qualitative and quantitative traits

The core set had higher qualitative trait values of the Nei diversity index and Shannon-Wiener diversity index than the entire set (Table 2). The core set was further validated by estimating the coefficient of variation, variance, and means for each quantitative character (Fig. 1). The fifteen quantitative traits in the core set had a higher coefficient of variance than their counterparts in the base population. In terms of variance, the core set outperformed the base set with the exception of the number of leaves, leaf width, and pseudo stem diameter, which showed little, or null variation in both sets. The core set's mean values of quantitative characters were found to be higher than the entire set's (Fig. 2), except for the number of cloves per bulb. In the entire and core sets, a few minor traits including pseudo stem length, leaf width, polar diameter, equatorial diameter, and pseudo stem diameter showed similar means with less variation.

3.4. Characterization of the core set accessions

3.4.1. Phenotypic characterization and cluster analysis

The traits showed significant variation among accessions, with a coefficient of variance ranging from 12.63 to 30.83, suggesting that the trial was well-run, the traits as total yield, number of cloves per bulb, marketable yield, equatorial diameter, polar diameters, and pseudo stems length, TSS, average bulb weight, leaf length, plant height, weight of 50 cloves, and days to maturity were the variables that varied the most.

According to the dendrogram (Fig. 3), the core set was predominantly divided into two groups, with 22 and 24 accessions in Groups I and II, respectively, and the greatest genetic distance was found between accession 650 and IC-141325, i.e. 9.99. There was no association between the accessions' native location and their grouping pattern, except for two accessions (IC-372921 and DOGR-593). These two accessions are native to India's Jammu and Kashmir region, with average bulb weights of less than five grams (low, or c-grade) and the lowest extremities for mean values in related traits. In addition, three more accessions, WG-418 (Andhra Pradesh), IC-372947 (Gujarat), and RG-343 (Maharashtra), provided marketable yields of less than 8 q/ha with 80% low-grade bulb.

In terms of plant architecture, four accessions had spreading leaf postures, 11 had erect postures, and the rest were semi-spreading. Six accessions matured in less than 120 days, with RG-95 becoming the first to mature (106–110 days). The total soluble solids (TSS) range of the set was 35.5 to 45.8°B, with the highest value being IC-35582 (Uttar Pradesh). The bulb's equatorial diameter ranged from 3.0 to 21.9 mm, with just three accessions exceeding 15 mm in diameter namely WG-101 (Rajasthan), 650 (Orissa), and IC-375028 (Gujarat), and in the core set's marketable yield varied between 6.0 and 70.0 q/ha. Cluster I and cluster II are represented in the current core range by accessions IC-35582, RG-95, and 543, which were chosen for their high TSS, early maturity, and highest marketable potential, respectively. These groups had genotypes with almost the same amount of unique traits and could be used to replace the above-mentioned accessions if they were lost.

Table 1
Different approaches of core set extraction and respective statistical indicators.

Method used	Non-heuristic approach	Random selection method
Number of entries processed as entire set	625	625
No of entries selected as core set	46	63
Mean Difference % (MD)	8.4	7.04
Variance Difference % (VD)	56.12	50.57
Confidence Ratio % (CR)	97.78	97.82
Variable Rate % (VR)	150.18	140.26
Efficiency Index	0.8	–

3.4.2. Biochemical characterization of core set accession and its cluster analysis

The TSS of the garlic core set ranged from 35 to 46 °B. Pyruvic acid (PA) levels in the core set ranged from 14.4 to 52.8 mol gm-1 FW (fresh weight), with accession 543 having the highest PA level, followed by 520, IC141325, and 549. Accession WG-432, No-15, and IC-48651, on the other hand, had the lowest PA value. The concentrations of phenol and protein are almost normally distributed, with significant differences between accessions. The highest phenol concentration was found in genotype 543 (901.72 mg/kg), while the lowest was found in accession 161 (321.75 mg/kg), and in case of proteins, accession 272 (3.1%) exhibited the highest (Fig. 3, Table S2). The total allicin content of the core set ranged from 1.78 to 7.01 mg/g (Table S2). Eight of the 46 core set accessions had higher allicin content than the British

Table 2
Different diversity indices and comparison between core set and entire set for qualitative traits.

Character	Core set collection			Entire set collection		
	Shannon-Wiener Diversity Index	Nei Diversity Index	C-Allele	Shannon-Wiener Diversity Index	Nei Diversity Index	E-Allele
Plant Architecture	0.707	0.351	3	0.34	0.148	3
Bulb Shape	1.134	0.656	5	0.932	0.587	5
Bulb Skin Colour	1.456	0.74	7	1.046	0.611	7
Clove Skin Colour	0.623	0.414	3	0.563	0.408	3

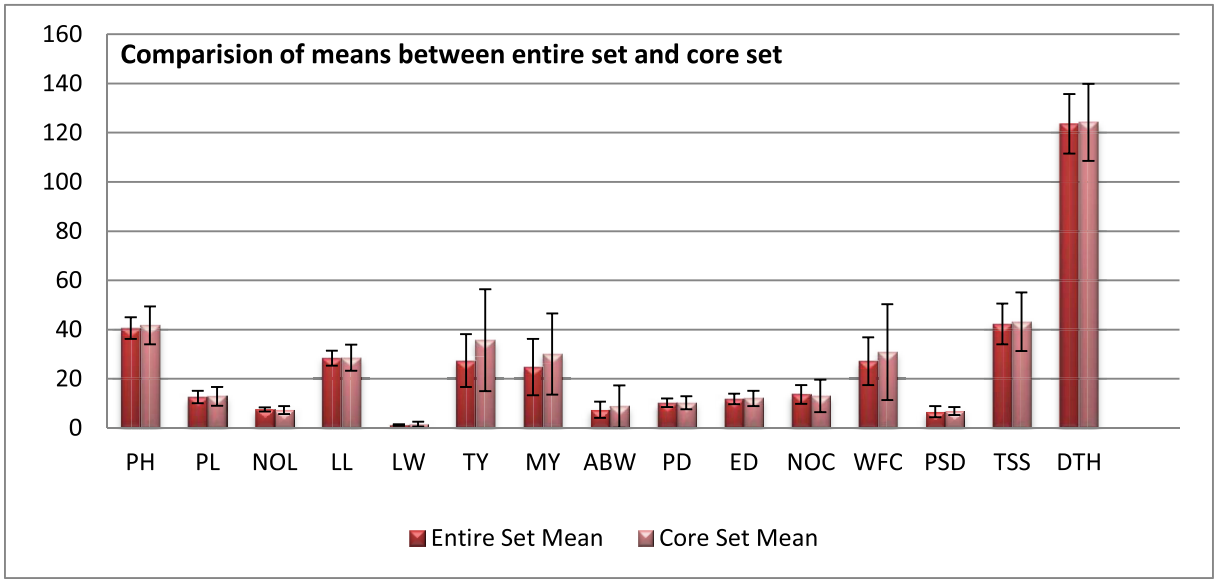

Fig. 2. Comparison between means of quantitative traits between entire set and core set.

herbal pharmacopeia record (4.5 mg/g). Four hard neck accessions (showing bolt and scape formation) indicated an allicin level below the pharmacopeia level, which was contrary to our findings.

Pair-wise genetic distance (GD) values across garlic core set genotypes ranged from 0 to 8, with an average GD of 2.22 across all genotypes. A pair-wise analysis of 16 accessions revealed GD less than one, suggesting genetically similar accessions for the traits studied. To anticipate the genetic relationships, the dendrogram was constructed using pair-wise genetic distance. The genotypes were divided into two large clusters, each with 1.74 genetic distance values. Additional accessions will be ordered according to their marketable yield capacity. Later, classifications based on bio compounds were added (Fig. 4). Within the core accessions analysed, there was a lot of variation in chemical content. Based on the information gathered, the entire core set was divided into two clusters, each with 29 and 17 accessions. The second category has the highest levels of allicin and complete soluble solids, as well as the highest levels of allyl methyl thiosulphinate (AMThs) and allyl trans-1-propenyl thiosulphinate (ATPThs). The first group was in favour of grouping accessions with high PA, however, since there was no association between geographical origin and PA content and the representation of bolted genotypes did not follow the clustering pattern (No-15, IC-34582, and IC-32881).

3.4.3. Core accessions characterization using SRAP markers and its cluster analysis

The 30 screened primer combinations deployed in core set characterization, amongst the primer combinations namely EM2-ME3, EM2-ME5, EM3-ME1, EM3-ME5, EM5-ME1, EM5-ME3, EM6-ME4, and EM6-ME5 had 100 per cent polymorphism, followed by EM1- ME3 (87.50 per cent), EM1-ME5 (88.89 per cent), and 85.71 per cent polymorphism with primer combinations EM1-ME2, EM2-ME4, EM4-ME1, and EM5-ME4, and the combination EM1-ME2 yielded the lowest polymorphism (33.33%). The total polymorphism was determined to be 80.59 per cent, on an average, the 30 SRAP primer combinations produced 5.1 bands per primer with 4.16 polymorphic and 1.0 monomorphic bands, the number of bands varies from 2 to 9 which ranges from 100 to 1500 base pairs. The primer combination EM6- ME2 had the highest polymorphism information content (PIC) of 0.55, while the primer combination EM5- ME1 and EM5-ME5 had the lowest (0.10). The average PIC of 30 primer combinations was 0.29 (Table 4). The primer combination EM1-ME5 noted the highest effective multiplex ratio (EMR) of 7.11, while the primer combination EM1-ME2 revealed the lowest EMR of 0.33. The total EMR average of 30 primer combinations was 3.51. The highest marker index (MI) was 2.34 with primer combination EM1-ME5, the lowest was 0.23 with primer combination EM4-ME2, and the average of 30 primer combinations was 0.99.

A dendrogram based on Jaccard's similarity revealed that 46 accessions were grouped into two major clusters, each with 40 and 6 accessions (Fig. 5). The major cluster I had six sub-clusters are IA (6 accessions), IB (7 accessions), IC (24 accessions), ID (1 accession), IE, and IF (2 accessions in each sub-cluster). Although major cluster II was divided into two sub-clusters, sub-cluster IIA (1 accession) and sub-cluster IIB (3 accession). At similarity coefficients of 0.75 and 0.61, respectively, and the accessions, No. 644 and IC-82882 did not form a cluster and remained isolated. The dendrogram revealed a near relationship between No. 644 and M−199, as well as RG-61 and IC-141249. As they fell on opposite sides of the dendrogram, the accessions IC-375119, IC-372924,

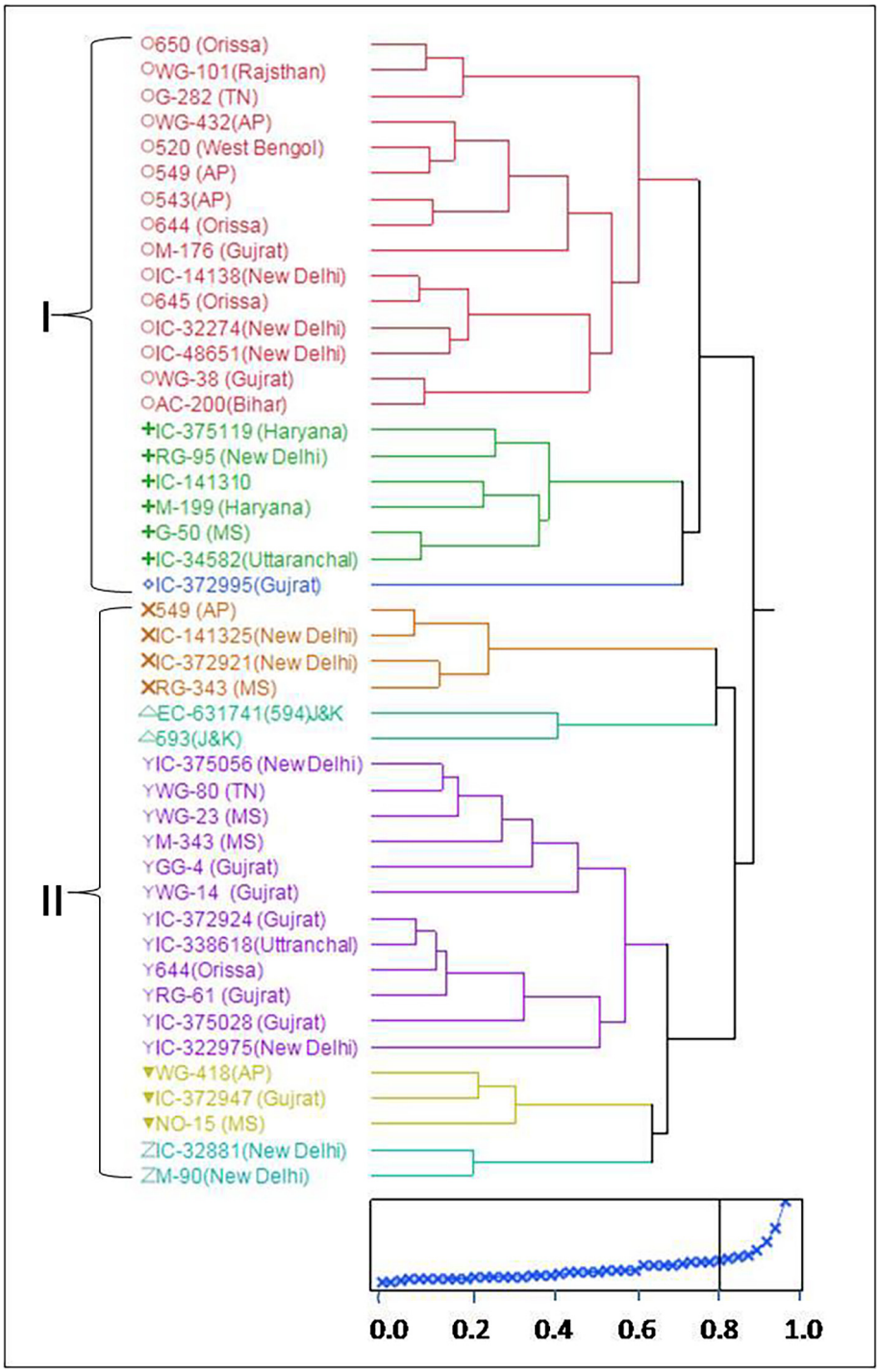

Fig. 3. Genetic distance based dendrogram of 46 garlic core set accessions based on Agro-morphological traits. Parenthesis in front of accessions name shows place of origin of respective genotypes. There no correlation observed between geographic origin and grouping of genotypes.

and IC-372953 were isolated from G-282 and IC-244959, suggesting a high genetic diversity among these five accessions.

4. Discussion

In clonal propagated crops, the state of redundancy and the chances of repetition are more, and garlic is no exception. The genetic resource for Indian garlic in ICAR-DOGR today is the product of twenty-five years of exploration tours. Although there are more garlic entries available in gene bank, more chance of duplication also present. As a result, the centre has agreed to limit the number by focusing on core set development. In the whole set, there was a non-significant association of genotypes years at p less than 0.01, p less than 0.05 for all traits. This may be due to garlic's clonally propagation process, which has been used for years, and its inherent sterility, which restricts exposure to its natural diversity (Benke et al., 2020b). The number of leaves was found to be important in all parameters except leaf distance, polar diameter (mm), equatorial diameter (mm), and pseudo stem diameter (mm). The coefficient of variation among various parameters

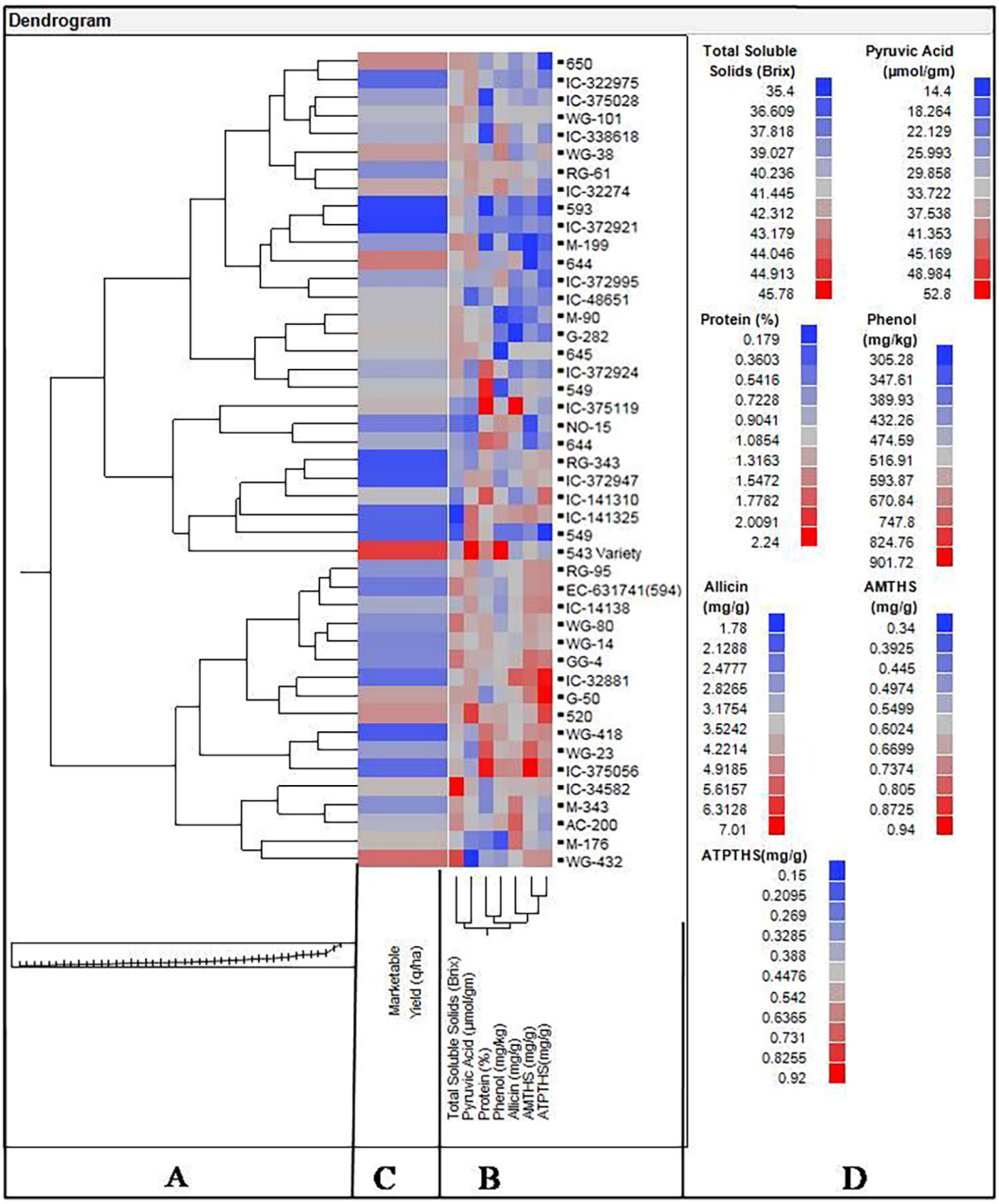

Fig. 4. (A) Dendrogram demonstrating genetic relationships among 46 core set garlic accessions. The dendrogram is based on Ward's Genetic distance method. Clusters 1–2 indicate the groups revealed by clustering the garlic accessions at GD < 12. (B)Total soluble solids, Allicin, AMTHS, ATPHS, Pyruvate, Protein and Phenolics (Red-Blue-Gray Scale) contents were sub-classified into quartiles and indicated by their colour intensity, with darker colour representing higher concentration, and lighter colours for lower concentration. (C) Estimated genetic structure of the garlic accessions further classified with respect to Marketable yield. Each accession is represented by a horizontal bar partitioned into three coloured segments depicting each individual. (D) Legends of traits with colour intensity scale and its respective numerical value.

had a low range value (9.82–30.39 percent), indicating that the trial was carried out with sufficient precision (Fig. 1). Several scientific studies have found genetic variation in a range of agro-morphological and biochemical traits in garlic. Core set production is one of the crop management strategies for maintaining genetic diversity and ensuring its successful exploitation (Panthee et al., 2006; Khar et al., 2011; Mostafa et al., 2015;Barboza et al., 2020; Benke et al., 2020a)

4.1. Characterization of agro-morphological traits of core set

The mean data of 18 morphological traits of 625 garlic accessions was analysed for core set development .The heuristic core selection (HCC) method sorted 46 accessions by an indiscriminate selection process into a core set, as the statistical indicators for core set validation are VR, VD, CR, and MD per cent (Kim et al., 2007; Gowda et al., 2012). As a result, we tested both core sets for the following conditions: (1) mean difference percentage 20%, i.e. less than 20% of characters displayed a different average between the base population and core collection, and (2) core set coincidence rate (CR per cent) should not be less than 80% of the characters noted (Hu et al 2000). By presenting the entire array, the CR obtained for both core sub-sets was greater than 80% with high VD and VR, the core set is recognized as reflecting higher genetic diversity. The higher VD value indicates that there was more variation among the traits found in the core set relative to

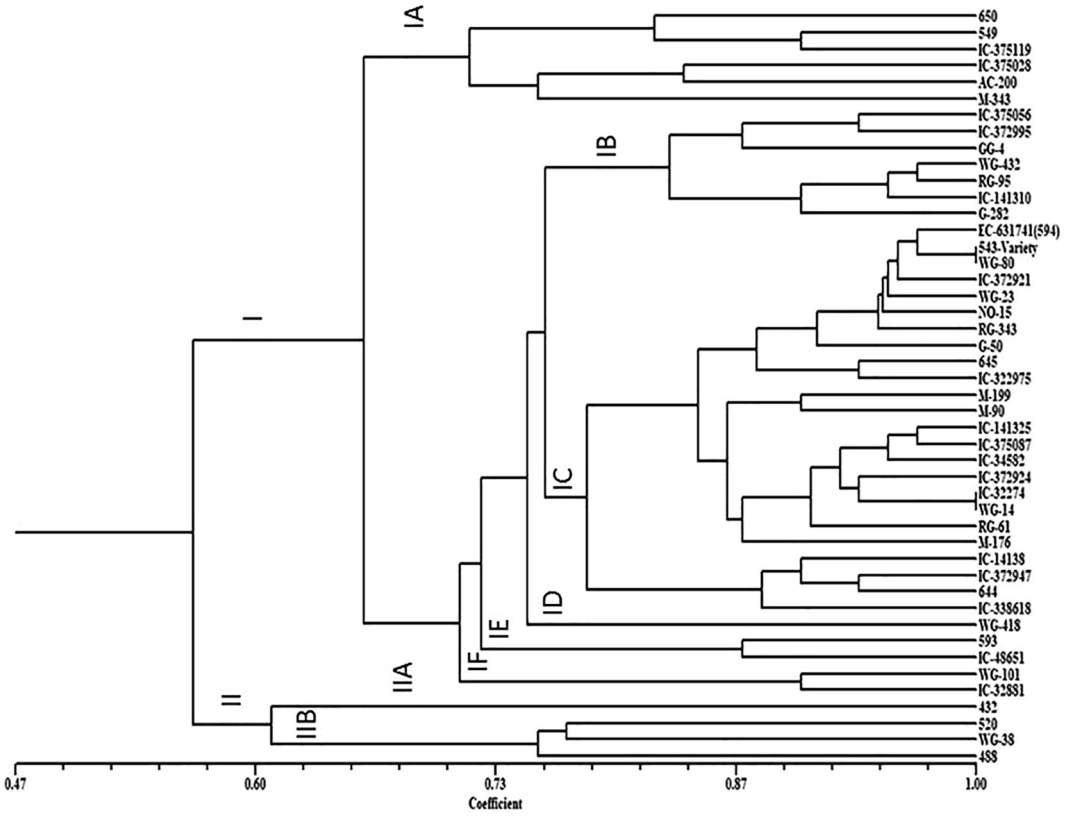

Fig. 5. UPGMA dendrogram based on the Nei's genetic similarity index illustrating the genetic relationship among 46 garlic genotypes using SRAP markers.

the base selection. However, a VR value greater than cent per cent indicates that the CV for all of the traits used in core set formulation was higher than the intact array. In comparison to the random process, the HCC approach sorted a smaller number of accessions with less redundancy and a higher VD per cent with more precision and performance (Table 1) (Kim et al., 2007). In addition, the HCC method provided 100% coverage for all of the evaluated characters. Since there are fewer accessions, it would be easier to handle, maintain, and examine some of the more nuanced traits. HCC was considered and a shorted group of 46 accessions was selected as the core set for future experiments against the random selection method since this sorted set satisfies our aim of formulating the core set. These findings back up the theory of rice (Chung et al., 2009), in foxtail millet (Gowda et al., 2012) and in wheat (Dutta et al., 2015). Further validated genetic diversity present in studied qualitative traits and their inclusion in the core set, the Shannon-Wiener diversity index (Shannon and Weaver, 1949) and the Nei diversity index (Nei, 1973) were used. When comparing the core set to the entire set, the qualitative trait values of the Nei diversity index and Shannon-Wiener diversity index were higher for the core set (Table S2). Higher or even comparable values are not considered to be indicative of a strong core selection package. Other researchers (Holbrook et al 1993; Mahalakshmi et al., 2007; Dutta et al., 2015) have found similar results to ours.

The coefficient of variation, variance, and means were estimated to further validate the core collection quantitative characters. The core set's coefficient of variance was higher than the base population's for all fifteen quantitative traits. In most cases, the core set surpassed the base set, with the exception of the number of leaves, leaf width, and pseudo stem diameter, where there was little to no difference between the two sets (Table 2). The high variance in plant height and TSS suggests that garlic may be used in salads (Chung et al., 2009), with a focus on leaf morphology, leaf chemistry, and leaf waxiness (Singh et al., 2020) for medicinal and nutraceutical purposes. The core set's mean values of quantitative characters were found to be higher than the entire set's, except for the number of cloves per bulb (Table 2). In the entire and core sets, a few minor traits including pseudo stem length, leaf width, polar diameter, equatorial diameter, and pseudo stem diameter showed similar means with less variation. In the case of foxtail millet, a similar pattern was observed for the above indices in eleven quantitative traits (Gowda et al., 2012). As a consequence, the new core set is an illustrative set of the entire selection, maintaining a high degree of genetic variation.

4.2. Cluster analysis of core set accessions based on agro-morphological traits

All accessions reflects diversity for a group of traits, it's fascinating to learn about the associations between the core set accessions. Cluster analysis revealed a lot of genetic variation depending on screened traits. According to the dendrogram, the core set was divided into two classes (Group I and II, which included 22 and 24 accessions, respectively), with the greatest genetic distance between accession No. 650 and IC-141325 (Fig. 3). Except for two accessions (IC-372921 and DOGR-593), which are native to Jammu and Kashmir, India, and had low average bulb weight with the lowest extremities for mean values in related traits, there was no association between the native place of accessions and their grouping pattern. These accessions require 14–16 h of photoperiod and the lowest day and night temperature range to complete their life cycle, rendering them unsuitable for short day climatic conditions. In terms of plant architecture, four accessions had spreading leaf postures, eleven had erect postures, and the rest had semi-spreading leaf postures. Six accessions matured in less than 120 days, with RG-95 becoming the first to mature (106–110 days).

The total soluble solids of IC-35582 (Uttar Pradesh) was noted largest, and the equatorial bulb diameter of accessions namely WG-101 (Rajasthan), 650 (Orissa), and IC-375028 (Gujarat) were all greater than 15 mm (Table S2). In the core set, marketable yield varied between 6.0 and 70.0 q/ha. Indian ecotypes have nearly half the yield of Chinese ecotypes, as well as lower levels of other associated traits like bulb diameter, bulb weight, and vegetative growth parameter (Chen et al., 2014). Due to similar climatic conditions, Indian garlic capacity appears to be strong or similar to Nepali garlic ecotypes (Panthee et al., 2006). Another application of core set formulation, according to Holbrook et al., 1993, is that in the absence or loss of any accession with trait-specific character, it can be easily substituted with adjacent accessions present in that cluster. Cluster I and cluster II are represented in the current core range by accessions IC-35582, RG-95, and 543, which were chosen for their high TSS, early maturity, and highest marketable potential, respectively. These groups had genotypes with about the same amount of unique traits and could be used to replace the above-mentioned accessions if they were lost.

4.3. Biochemical characterization of core set accession

The TSS of the current garlic core set ranged from 35 to 46 °B, which is close to Greek garlic ecotypes (Petropoulous et al., 2018) and follows another composite set of garlic from India (Bhusal et al., 2019). However, accumulation of high TSS is not limited to the origin of genotypes in the current collection, which is in contrast, who observed clustering of Northern ecotypes with high TSS (Bhusal et al., 2019). The total soluble solid reflects the garlic cloves total sugar content. TSS is also linked to the long-term survival of dehydrated goods and the bulb's storage life (Barboza et al., 2020).

In the core set, PA ranged from 14.4 to 52.8 μmol gm/FW (fresh weight) and accession 543 recorded the highest PA followed by 520, IC141325, and 549. Accession WG-432, No-15, and IC-48651, on the other hand, had the lowest PA value. Low pungent cultivars can be useful in salad preparation, while accessions with high PA content would be more noticeable in medicinal uses such as lowering total blood cholesterol (Bhusal et al., 2019; Natale et al., 2004). The presence of pyruvic acid (PA) is usually linked to the characteristic garlic odour. Its pungency is derived from S-alk (en)yl-L-cysteine sulphoxide, as well as other volatile and non-volatile compounds, in the presence of the enzyme allinase (Khar et al., 2011). The Indian garlic accessions were classified as low pungency when they contained less than 60 mol/g PA, according to Natale et al.(2004). However, Bhusal et al.(2019) registered a PA range of 49.67 to 76.35 mol/g PA for an Indian composite set of varieties and landraces. This indicates why Indian pungent ecotypes are in such high demand in the pharmaceutical industry.

The content of phenol (TPC) and protein in the core set almost normally distributed, with major variations between accessions. The highest phenol concentration was found in accession No.543, while the lowest was found in accession No.161, TPC variation was documented in previous studies of garlic, including Chinese ecotypes 42.53 mg GAE/g (Mostafa et al., 2015), 223–394 g/g in genotypes from Poland (Bozin et al., 2008), 0.98 mg GAE/g from Serbia (Kim et al., 2013).

In case of proteins, accession no.272 exhibited the highest value, the obtained findings are inconsistent with a previously published study by grouping garlic accessions into different classes, suggesting that the amount of phenol and protein varied in different garlic accessions (Szychowski et al., 2018; Kim et al., 2013), and that it indirectly depends on clove (Beato et al., 2011). As a result of these findings, Indian garlic ecotypes seem to have a bright future as a supplement and pharmaceutical product

The content of allicin and other thiosulphinates differed significantly among core accessions, according to the results of this experiment. The total allicin content of the 46 accessions ranged from 1.78 to 7.01 mg/g (Table S2). *A sativum* L. has the highest thiosulphinate content of all the Alliums (Singh et al., 2020). The amount of allicin found in the garlic studied set is almost identical to that found in Egyptian (Mostafa et al., 2015) and Argentinean (Natale et al., 2004) garlic. This major variation in ecotypes based on agro morphological and chemical content corresponds to a previous study on garlic thiosulphinates (Khar et al., 2011; Mostafa et al., 2015; Singh et al., 2020).

Among core set accessions, eight accessions in core set had higher allicin content (4.5 mg/g) than the British herbal pharmacopeia record. This revealed that the total number of accessions (with high pharmaceutical value) is higher than in the previous study by Khar et al., (2011). This explains why and how a natural and core collection of garlic ecotypes should be screened, and in other Alliums (Singh et al., 2020). Clonal selection can be used to increase cloves of those accessions with high allicin yield (Mostafa et al., 2015). These accessions will also be useful as a source of garlic supplement development content. The hard-neck garlic (grown in temperate region) gives higher potential allicin than soft-neck garlic (grown in tropical and subtropical region) (Gonzalez et al., 2009). Four hard neck accessions (showing bolt and scape formation) indicated an allicin level below the pharmacopeia level, which was contrary to our findings, the allicin content of Indian germplasm (Khar et al., 2011) is comparable to that of Egypt and China (Mostafa et al., 2015). Geographic location, biotic and abiotic conditions and the system used for analysis are influences allicin levels (Panthee et al., 2006; Gonzalez et al., 2009; Barboza et al., 2020). As a result, allicin content ranged across genotypes, progeny, and location in previous studies.

4.4. Diversity analysis based on biochemical traits

To predict genetic relationships, the dendrogram was created using pair-wise genetic distance. The genotypes were grouped into two main clusters, each with 1.74 genetic distance values (Fig. 1). Pair-wise genetic distance (GD) values among garlic core set genotypes ranged from 0 to 8, with an average GD of 2.22 across all genotypes. The accessions were compared pair-wise, GD was less than one (16 accessions), suggesting that the accessions were genetically identical for the traits studied, whereas other accession in the core set, there was a lot of variation in chemical content. Among two clusters, the genotypes in the second cluster contain the highest levels of allicin and total soluble solid, as well as Allyl methyl thiosulphinate (AMThs) and Allyl *trans*-1-propenyl thiosulphinate (ATPThs). The first cluster, on the other hand, sorted accessions with high PA. Since there was no association between geographical origin and PA content (Gonzalez et al., 2009). Simillar results obtained by other researchers (Khar et al., 2011,; Barboza et al., 2020; Singh et al., 2020)

4.5. Characterization using SRAP markers

Thirty primer combinations showed polymorphism ranging from 33.33 to 100 per cent, with an overall polymorphism of 80.59 percent. Primer combinations generated an average of 5.1 bands per primer, with 4.16 polymorphic and 1.0 monomorphic bands, the number of bands ranges from 2 to 9, with 100 to 1500 base pairs between them. The primer pair EM6 ME2 had the highest PIC (0.55), the primer pair EM1 ME5 had the highest EMR (7.11), and the primer pair EM1 ME5 had the highest MI (2.34) (Table 3). For genetic diversity studies in garlic accessions, similar results were confirmed using EST, SSRs (Zhao et al., 2011; Chand

Table 3
Basic statistics of 15 quantitative Agro-morphological traits of core set.

Traits	Range	Mean	Standard Error	Standard deviation	Coefficient of variance
Plant Height (cm)	46.31	41.66	1.14	7.67	18.41
Pseudo stem length (cm)	19.44	12.82	0.56	3.76	23.35
Number of leaves	8.8	7.31	0.24	1.62	22.18
Leaf Length (cm)	27.92	28.58	0.79	5.31	18.59
Leaf Width (cm)	4.61	1.56	0.15	0.98	13.92
Total Yield (q/ha)	162.41	35.67	3.98	26.69	30.83
Marketable yield (q/ha)	68.18	30.05	2.47	16.54	28.05
Average Bulb Weight (gm)	59.48	8.69	1.29	8.65	18.11
Polar Diameter (mm)	13.63	10.27	0.39	2.65	25.81
Equatorial Diameter (mm)	18.9	11.99	0.47	3.14	26.17
Number of Cloves/Bulb	38.35	13.02	0.98	6.56	30.39
Weight of 50 Cloves (gm)	111.5	30.78	2.91	19.51	16.38
Pseudo Stem Diameter (mm)	11.2	6.88	0.25	1.65	21.98
Total Soluble Solids (°Brix)	84.6	43.18	1.77	11.87	22.49
Days to Harvest	133.78	124.23	2.34	15.67	12.62

Table 4
Detail of SRAP primers combinations and their amplification profile used in genetic diversity assessment.

S.N.	Primers	Tm (°C)	AB	SR (bp)	MB	PB	PE	PIC	EMR	MI
1	EM 1 ME 1	49.75	6	250–1500	1	5	83.33	0.26	4.17	1.10
2	EM 2 ME 1	52.05	5	250–1500	1	4	80.00	0.15	3.20	0.47
3	EM 3 ME 1	52.05	4	250–1500	0	4	100.00	0.22	4.00	0.90
4	EM 4 ME 1	50.90	7	250–1500	1	6	85.71	0.35	5.14	1.78
5	EM 5 ME 1	50.40	5	250–1500	0	5	100.00	0.10	5.00	0.51
6	EM 6 ME 1	52.05	3	200–500	1	2	66.67	0.18	1.33	0.24
7	EM 1 ME 2	52.15	7	220–1000	1	6	85.71	0.21	5.14	1.05
8	EM 2 ME 2	54.45	3	100–300	2	1	33.33	0.33	0.33	0.11
9	EM 3 ME 2	54.45	6	200–990	2	4	66.67	0.29	2.67	0.77
10	EM 4 ME 2	53.30	2	190–470	1	1	50.00	0.46	0.50	0.23
11	EM 5 ME 2	55.20	4	250–1500	1	3	75.00	0.22	2.25	0.50
12	EM 6 ME 2	54.45	6	250–1500	1	5	83.33	0.55	4.17	2.31
13	EM 1 ME 3	49.75	8	100–700	1	7	87.50	0.22	6.13	1.35
14	EM 2 ME 3	52.00	3	250–1500	0	3	100.00	0.20	3.00	0.59
15	EM 3 ME 3	52.00	4	200–700	2	2	50.00	0.26	1.00	0.26
16	EM 4 ME 3	50.90	5	310–680	2	3	60.00	0.41	1.80	0.73
17	EM 5 ME 3	50.40	2	200–300	0	2	100.00	0.36	2.00	0.72
18	EM 6 ME 3	52.00	4	130–700	1	3	75.00	0.26	2.25	0.58
19	EM 1 ME 4	52.10	6	190–1500	1	5	83.33	0.36	4.17	1.49
20	EM 2 ME 4	54.45	7	250–3000	1	6	85.71	0.33	5.14	1.68
21	EM 3 ME 4	54.45	6	270–5000	2	4	66.67	0.48	2.67	1.27
22	EM 4 ME 4	53.30	4	450–1000	1	3	75.00	0.45	2.25	1.01
23	EM 5 ME 4	55.20	7	250–1500	1	6	85.71	0.28	5.14	1.43
24	EM 6 ME 4	54.45	4	250–1500	2	4	100.00	0.21	4.00	0.82
25	EM 1 ME 5	50.95	9	250–1500	1	8	88.89	0.33	7.11	2.34
26	EM 2 ME 5	53.25	6	250–1500	0	6	100.00	0.26	6.00	1.54
27	EM 3 ME 5	53.25	3	100–310	0	3	100.00	0.36	3.00	1.07
28	EM 4 ME 5	52.10	4	200–1500	1	3	75.00	0.35	2.25	0.78
29	EM 5 ME 5	52.80	8	180–700	2	6	75.00	0.10	4.50	0.46
30	EM 6 ME 5	53.25	5	250–1500	0	5	100.00	0.34	5.00	1.71
Average Performance			**5.1**	**–**	**1**	**4.16**	**80.59**	**0.29**	**3.51**	**1.00**

TA, Annealing temperature; AB, Amplified bands; SR, Size range base pair; PB, Polymorphic bands; MB, Monomorphic bands; PE, Polymorphism efficiency; PIC, Polymorphism information content; EMR, Effective multiplex ratio; MI, Marker index.

et al., 2015; Kumar et al., 2019; Barboza et al., 2020) and SRAP markers (Li et al., 2007).

4.6. Cluster analysis using SRAP markers

Dendrogram based on Jaccard's similarity core set divided into two major clusters, major cluster I divided into six sub-clusters, IA (6 accessions), IB (7 accessions), IC (24 accessions), ID (1 accession), IE, and IF (2 accessions in each sub-cluster). While major cluster II was separates in two sub-clusters, sub-cluster IIA (2 accession) and sub-cluster IIB (4 accessions). The accessions No.650 and M−199, as well as RG-61 and IC-141249, were very close similarity. The accessions IC-375119, IC-372924, and IC-372953 were isolated from G-282 and IC-244959, as they fell on opposite sides of the dendrogram, suggesting a high genetic diversity among these five accessions. The clustering of garlic accessions based on SRAP markers data did not match the clustering based on biochemical traits, but the result was consistent with (Li et al., 2007). The current finding indicated that the garlic core set produced had a lot of genetic diversity. A genetically distinct collection of garlic accessions may help expand the genetic base and give strength to the current breeding program.

5. Conclusion

Germplasm maintenance in the field entails special attention, and environmental constraints such as biotic and abiotic stresses may result the permanent loss of genotypes. As a result, this work was carried out to make the study and maintenance of Indian garlic germplasm easier. The current core collection research establishes

primary data about Indian garlic collection based on defined traits, allowing crucial evaluation of various traits more accessible through an interdisciplinary approach. Almost all garlic growing status in India is reflected by this core group. The core represents 18 agro-morphological traits, with almost maximum variability of genotypes, additional assessment will enable the core selection to be refined. More specifically, the current research would aid in the preservation of genetically diverse Indian garlic germplasm.

Declaration of Competing Interest

None.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.sjbs.2021.05.013.

References

Abdoli, M., Habibi, K.B., Baghalian, K., Shahnazi, S., Rasouli, H., Naghdi, B.H., 2009. Classification of Iranian garlic (Allium sativum L.) ecotypes using RAPD marker., 45-51.

Abedi, M., Biat, F., Nosrati, A.E., 2013. Evaluation of agronomical traits and pyruvic acid content in Hamedan garlic (*Allium sativum* L.) ecotypes. Euro J. Expt Bio 3 (2), 541–544.

Barboza, K., Salinas, M.C., Acuña, C.V., Bannoud, F., Beretta, V., García-Lampasona, S., Burba, J.L., Galmarini, C.R., Cavagnaro, P.F., 2020. Assessment of genetic diversity and population structure in a garlic (*Allium sativum* L.) germplasm collection varying in bulb content of pyruvate, phenolics, and solids. Sci Hort 261, 108900.

Beato, V.M., Orgaz, F., Mansilla, F., Montaño, A., 2011. Changes in phenolic compounds in garlic (*Allium sativum* L.) owing to the cultivar and location of growth. Plant food human nutri 66 (3), 218–223.

Benke, A.P., Dukare, S., Jayaswall, K., Yadav, V.K., Singh, M., 2019. Determination of proper gamma radiation dose for creating variation in Indian garlic varieties. Ind J. Trad Know 18 (3), 547–552.

Benke, A.P., Khar, A., Mahajan, V., Gupta, A., Singh, M., 2020a. Study on dispersion of genetic variation among Indian garlic ecotypes using agro morphological traits. Ind J. Genet. 80 (1), 94–102.

Benke, A.P., Nair, A., Krishna, R., Anandhan, S., Mahajan, V., Singh, M., 2020b. Molecular screening of Indian garlic genotypes *(Allium sativum* L.) for bolting using DNA based Bltm markers. Veg. Sci. 47 (1), 116–120.

Bhusal, R., Islam, S., Khar, A., Singh, S., Jain, N., Tomar, B.S., 2019. Diversity analysis and trait association study for antioxidants and quality traits in landraces, farmers' varieties and commercial varieties of Indian short day garlic (*Allium sativum* L.). Genet Resour Crop Ev 66 (8), 1843–1859.

Bozin, B., Mimica-Dukic, N., Samojlik, I., Goran, A., Igic, R., 2008. Phenolics as antioxidants in garlic (*Allium sativum* L., Alliaceae). Food chem 111 (4), 925–929.

Chand, S.K., Nanda, S., Rout, E., Joshi, R.K., 2015. Mining, characterization and validation of EST derived microsatellites from the transcriptome database of *Allium sativum* L. Bioinformation 11 (3), 145.

Chen, S., Chen, W., Shen, X., Yang, Y., Qi, F., Liu, Y., Meng, H., 2014. Analysis of the genetic diversity of garlic (*Allium sativum* L.) by simple sequence repeat and inter simple sequence repeat analysis and agro-morphological traits. Biochem System Ecol 55, 260–267.

Chen, S., Shen, X., Cheng, S., Li, P., Du, J., Chang, Y., Meng, H., 2013. Evaluation of garlic cultivars for polyphenolic content and antioxidant properties. PLoS One 8, (11) e79730.

Chung, H.K., Kim, K.W., Chung, J.W., Lee, J.R., Lee, S.Y., Dixit, A., Gwag, J.G., 2009. Development of a core set from a large rice collection using a modified heuristic algorithm to retain maximum diversity. J. Integ Pl Bio 51 (12), 1116–1125.

Cunha, C.P., Hoogerhide Eulalia, S.S., Zucchi, M.I., Monterio, M., Pinheiro, J.B., 2012. New Microsatellite markers for garlic, *Allium sativum* (Alliaceae) American. J. Bot, e17–e19.

Dutta, M., Phogat, B.S., Kumar, S., Kumar, N., Kumari, J., Pandey, A.C., Singh, T.P., Tyagi, R.K., Jacob, S.R., Srinivasan, K., Bisht, I.S., 2015. Development of core set of wheat (*Triticum* spp.) germplasm conserved in the national genebank in India. In Advances in wheat genetics: from genome to field. Springer, Tokyo, pp. 33–45.

Etoh, T., Simon, P.W., 2002. Diversity, fertility, and seed production of garlic. In: Rabinowvitch, H.D., Currah, L. (Eds.), *Allium* crop science: Recent advances. CABI Publishing, Wallingford, UK, pp. 101–117.

Folin, O., Denis, W.A., 1915. colorimetric method for the determination of phenols (and phenol derivatives) in urine. J of Bio Chem 22 (2), 305–308.

Gangopadhyay, K.K., Mahajan, R.K., Kumar, G., Yadav, S.K., Meena, B.L., Pandey, C., Bisht, I.S., Mishra, S.K., Sivaraj, N., Gambhir, R., Sharma, S.K., 2010. Development of a core set in brinjal *(Solanum melongena* L.). Crop sci 50 (3), 755–762.

Gonzalez, R.E., Soto, V.C., Sance, M.M., Camargo, A.B., Galmarini, C.R., 2009. Variability of solids, organosulfur compounds, pungency and health-enhancing traits in garlic (*Allium sativum* L.) cultivars belonging to different ecophysiological groups. J. of agri and food chem 57 (21), 10282–10288.

Gowda, J., Krishanappa, M., Pathak, N., Mathur, P.N., Seetharam, A., 2012. Use of heuristic approach for the development of a core set from large germplasm collection of foxtail millet (*Setaria italica* L.). Ind J. Pl Genet Resour 26 (1), 13–18.

He, F., 2011. Bradford protein assay. Bio Protoc. 101, e45.

Holbrook, C.C., Anderson, W.F., Pittman, R.N., 1993. Selection of a core collection from the US germplasm collection of peanut. Crop sci, 33(4), pp.859-861 http://www.fao.org/faostat/en/#data/QC. FAOSTAT, 2018 http:/www.plantauthority.in (DUS) test guidelines of National Bureau of Plant Genetic Resources (NBPGR) in New Delhi, India

Hu, J., Zhu, J., Xu, H.M., 2000. Methods of constructing core collections by stepwise clustering with three sampling strategies based on the genotypic values of crops. Theor. Appl. Genet. 101, 264–268.

Hua, G.J., Hung, C.L., Lin, C.Y., Wu, F.C., Chan, Y.W., Tang, C.Y., 2017. MGUPGMA: a fast UPGMA algorithm with multiple graphics processing units using NCCL. Ev Bioinform, 13, p.1176934317734220

Khandagale, K., Krishna, R., Roylawar, P., Ade, A.B., Benke, A., Shinde, B., Rai, A., 2020. Omics approaches in *Allium* research: Progress and way ahead. PeerJ 8, e9824.

Khar, A., Banerjee, K., Jadhav, M.R., Lawande, K.E., 2011. Evaluation of garlic ecotypes for allicin and other allyl thiosulphinates. Food Chem 128 (4), 988–996.

Kim, J.S., Kang, O.J., Gweon, O.C., 2013. Comparison of phenolic acids and flavonoids in black garlic at different thermal processing steps. J Fun Food. 5 (1), 80–86.

Kim, K.W., Chung, H.K., Cho, G.T., Ma, K.H., Chandrabalan, D., Gwag, J.G., Park, Y.J., 2007. PowerCore: a program applying the advanced M strategy with a heuristic search for establishing core sets. Bioinformatics 23 (16), 2155–2162.

Kumar, M., Sharma, V.R., Kumar, V., Sirohi, U., Chaudhary, V., Sharma, S., Sharma, S., 2019. Genetic diversity and population structure analysis of Indian garlic (*Allium sativum* L.) collection using SSR markers. Physio mol bio plants 25 (2), 377–386.

Li, H.Z., Yin, Y.P., Zhang, C.Q., Zhang, M., Li, J.M., 2007. Evaluation of Application of SRAP on Analysis of Genetic Diversity in Cultivars of Allium fistulosum L. Acta Horticulturae Sinica 34 (4), 929.

Mahalakshmi, V., Ng, Q., Atalobhor, J., Ogunsola, D., Lawson, M., Ortiz, R., 2007. Development of a West African yam Dioscorea spp. core collection. Genet Resour Crop Ev 54 (8), 1817–1825.

Manjunathagowda, D.C., Gopal, J., Archana, R., Asiya, K.R., 2017. Virus-Free Seed Production of Garlic (*Allium sativum* L.): Status and Prospects. Int. J. Curr. Microbiol. App. Sci. 6 (6), 2446–2456.

Mohana, G.S., Nayak, M.G., 2018. Development of the core collection through advanced maximization strategy with heuristic approach in cashew (Anacardium occidentale L.). Plant Genet Resour 16 (4), 367–377.

Morales, R., Resende, J.T.V., Resende, F.V., Delatorre, C.A., Figueiredo, A.S.T., Da-Silva, P.R., 2013. Genetic divergence among Brazilian garlic cultivars based on morphological characters and AFLP markers. Embrapa Hortaliças-Artigo em periódico indexado (ALICE).

Mostafa, H.H., Wang, H., Liu, X., Li, X., 2015. Impact of genetic factor and geographical location on allicin content of garlic (*Allium sativum*) germplasm from Egypt and China. Inter J of Agri and Bio 17 (1).

Natale, P.J., Camargo, A., Galmarini, C.R., 2004. Characterization of Argentine garlic cultivars by their pungency. IV Inter Sympo on Edi Alliaceae 688, 313–316.

Nei, M., 1973. Analysis of gene diversity in subdivided populations. Proceedings of the National Acad of Sci 70 (12), 3321–3323.

Oliveira, M.F., Nelson, R.L., Geraldi, I.O., Cruz, C.D., de Toledo, J.F.F., 2010. Establishing a soybean germplasm core collection. Field Crops Res 119 (2–3), 277–289.

Panthee, D.R., Kc, R.B., Regmi, H.N., Subedi, P.P., Bhattarai, S., Dhakal, J., 2006. Diversity analysis of garlic (*Allium sativum* L.) germplasms available in Nepal based on morphological characters. Genet Resource and Crop Evol 53 (1), 205–212.

Robarts, D.W., Wolfe, A.D., 2014. Sequence-related amplified polymorphism (SRAP) markers: A potential resource for studies in plant molecular biology1. Appl plant sci 2 (7), 1400017.

Schwimmer, S., Weston, W.J., 1961. Onion flavor and odor, enzymatic development of pyruvic acid in onion as a measure of pungency. J Agri Food Chem 9 (4), 301–304.

Shannon, C.E., Weaver, W., 1949. The mathematical theory of communication. University of Illinois, Urbana, p. 117.

Singh, P., Mahajan, V., TP, A.S., Banerjee, K., Jadhav, M.R., Kumar, P., Gopal, J., 2020. Comparative evaluation of different Allium accessions for allicin and other allyl thiosulphinates. Ind Crop Prod, 147, 112215

Singleton, V.L., Rossi, J.A., 1965. Colorimetry of total phenolics with phosphomolybdic-phosphotungstic acid reagents. Ameri J Eno Viti 16 (3), 144–158.

Szychowski, K.A., Rybczyńska-Tkaczyk, K., Gaweł-Bęben, K., Świeca, M., Karaś, M., Jakubczyk, A., Matysiak, M., Binduga, U.E., Gmiński, J., 2018. Characterization of active compounds of different garlic (*Allium sativum* L.) cultivars. Polish J. Food Nutri Sci 68 (1), 73–81.

Upadhyaya, H.D., Ortiz, R., Bramel, P.J., Singh, S., 2003. Development of a groundnut core collection using taxonomical, geographical and morphological descriptors. Genetic Resour Crop Ev 50 (2), 139–148.
Upadhyaya, H.D., Reddy, L.J., Gowda, C.L.L., Reddy, K.N., Singh, S., 2006. Development of a mini core subset for enhanced and diversified utilization of pigeonpea germplasm resources. Crop Sci 46 (5), 2127–2132.
Vaijayanthi, P.V., Ramesh, S., Gowda, M.B., Rao, A.M., Gowda, J., Ramappa, H.K., Keerthi, C.M., Rajendra Prasad, B.S., 2015. Development and validation of a core set of dolichos bean germplasm. Inter J Veg Sci 21 (5), 419–428.
Vavilov, N.I., 1951. The origin, variation, immunity and breeding of cultivated plants. LWW,72(6), 482.
Ward, J.H., 1963. Hierarchical grouping to optimize an objective function. J Am Stat Assoc. 58, 236–244.
Zhao, W.G., Chung, J.W., Lee, G.A., Ma, K.H., Kim, H.H., Kim, K.T., Chng, I.M., Lee, J.K., Kim, N.S., Kim, S.M., Park, Y.J., 2011. Molecular genetic diversity and population structure of a selected core set in garlic and its relatives using novel SSR markers. Pl. Breed 130, 46–54.
FAOSTAT, 2018 http://www.fao.org/faostat/en/#data/QC.