

Research paper

SSR markers development and their application in genetic diversity evaluation of garlic (*Allium sativum*) germplasm

Xiaxia Li, Lijun Qiao, Birong Chen, Yujie Zheng, Chengchen Zhi, Siyu Zhang, Yupeng Pan **, Zhihui Cheng *

College of Horticulture, Northwest A&F University, Yangling, 712100, Shaanxi Province, China

ARTICLE INFO

Article history:

Received 23 March 2021

Received in revised form

26 July 2021

Accepted 1 August 2021

Available online 11 August 2021

Keywords:

Garlic

SSR markers

Genetic diversity

Population structure

ABSTRACT

Garlic (*Allium sativum*), an asexually propagated vegetable and medicinal crop, has abundant genetic variation. Genetic diversity evaluation based on molecular markers has apparent advantages since their genomic abundance, environment insensitivity, and non-tissue specific features. However, the limited number of available DNA markers, especially SSR markers, are insufficient to conduct related genetic diversity assessment studies in garlic. In this study, 4372 EST-SSR markers were newly developed, and 12 polymorphic markers together with other 17 garlic SSR markers were used to assess the genetic diversity and population structure of 127 garlic accessions. The averaged polymorphism information content (PIC) of these 29 SSR markers was 0.36, ranging from 0.22 to 0.49. Seventy-nine polymorphic loci were detected among these accessions, with an average of 3.48 polymorphic loci per SSR. Both the clustering analyses based on either the genotype data of SSR markers or the phenotypic data of morphological traits obtained genetic distance divided the 127 garlic accessions into three clusters. Moreover, the Mantel test showed that genetic distance had no significant correlations with geographic distance, and weak correlations were found between genetic distance and the phenotypic traits. AMOVA analysis showed that the main genetic variation of this garlic germplasm collection existed in the within-population or cluster. Results of this study will be of great value for the genetic/breeding studies in garlic and enhance the utilization of these garlic germplasms.

Copyright © 2021 Kunming Institute of Botany, Chinese Academy of Sciences. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Garlic (*Allium sativum* L.), originated in central Asia, is one of the most worldwide cultivated and consumed horticultural crops since the ancient Egyptian period due to its edible and medicinal value (Vavilov, 1951; Hong and Etoh, 1996; Rahman and Lowe, 2006). Although garlic is asexually propagated, it shows surprisingly high biodiversity, environmental adaptation capacity, and phenotypic plasticity (Volk et al., 2004). A large-scale diversity of different cultivars has been established in various cultivation areas (Bradley et al., 1996; Avato et al., 1998; Baghalian et al., 2005; Wang et al., 2014). These abundant garlic cultivars or germplasms can

broaden the genetic variability and provide considerable opportunities for garlic genetics and breeding research (Zhao et al., 2011). The raw collections of different garlic cultivars, accumulated over the years, are usually classified based on phenotypic traits, which can easily lead to homonymy (same name for genetically different cultivars) and duplications or synonymy (same cultivars with different names) of garlic. That is especially problematic for the garlic with a similar appearance and significant phenotypic plasticity. Thus, how efficiently identify and distinguish of each garlic germplasms is of paramount importance to manage and maintain such genetic resources (Egea et al., 2017; Govindaraj et al., 2015).

The early identifications of garlic germplasms were mainly based on morphological characteristics that highly dependent on the field conditions and the local environments of their planting areas (Bradley et al., 1996; Al-Zahim et al., 1997). The genomic abundance, environment insensitivity, and non-tissue specific DNA markers showed significant advantages in biodiversity

* Corresponding author.

** Corresponding author.

E-mail addresses: yupeng.pan@nwafu.edu.cn (Y. Pan), chengzh@nwafu.edu.cn (Z. Cheng).

Peer review under responsibility of Editorial Office of Plant Diversity.

analysis of garlic and were gradually used in the identification and genetic diversity assessment of garlic accessions (Ovesná et al., 2014; Ipek et al., 2015). Furthermore, some DNA markers may be useful for marker-assisted selection in garlic breeding programs. Barboza et al. (2020) found markers AsESSR-30 and AsESSR-83 associated with flowering behavior, ecophysiological groups, and color types. The following types of markers including random amplified polymorphic DNA (RAPD) (Maaß and Klaas, 1995), amplified fragment length polymorphism (AFLP) (Ipek et al., 2005), simple sequence repeats (SSR) (Cunha et al., 2014), and insertions-deletions (InDel) (Wang et al., 2016) were all successfully used in garlic germplasm evaluations. Among these DNA markers, SSR markers became the preferred ones since they are typically codominant, reproducible, cross-species transferable, and highly polymorphic. SSR markers can be classified as genomic (gSSR) and genic (EST-SSR) SSRs, which were developed with the sequences of genomic DNA and cDNA or expressed sequence tags (ESTs), respectively (Vaek et al., 2020). Compared with the gSSR, the development of EST-SSR markers is relatively inexpensive because the required sequence data of cDNA or EST are easily obtained from public databases (Liu et al., 2015). Besides, primers of EST-SSRs are designed from more conserved coding regions of the genome, which made the EST-SSR markers more useful and with higher cross-species transferability (Varshney et al., 2005). However, the EST-SSRs are usually less polymorphic than gSSRs, and their PCR amplified fragments may fluctuate as the presence of introns in flanking regions (Kalia et al., 2011), which may lead to that not all EST-SSRs are perfect SSR markers. Therefore, the development of sufficient EST-SSRs is the precondition for fully using this kind of DNA markers. Recently, large numbers of EST-SSR markers have been successfully developed for many plant species (Lu et al., 2013; Blair and Hurtado, 2013; Mohanty et al., 2013). However, due to the relatively large genome size (~16 Gb) and the few available fertile germplasms, the molecular and genetic studies of garlic were largely lagged than other vegetable crops (Meryem et al., 2015). In addition, the available number of SSR markers is also relatively limited in garlic (Cunha et al., 2012). Previously, fewer than 100 SSR markers were reported in garlic (Ma et al., 2009; Lee et al., 2011; Cunha et al., 2012). Until recently, Liu et al. (2015) developed 1506 SSR markers using EST (expressed sequence tags) derived from a transcriptome dataset. Considering the relatively large genome size of garlic, the available SSR markers are still insufficient for adequate genetic studies in garlic (Liu et al., 2015).

Overall, using the SSR markers to genotype garlic cultivars from the molecular DNA levels will be an efficient way for the identification of each garlic germplasms since the environment insensitivity and non-tissue specific advantages of DNA markers. However, the shortage of sufficiently high-quality SSR markers is a serious limitation for the DNA marker-based identification of garlics. Therefore, the first objective of this study is to develop more new SSR markers in garlic based on the EST sequences derived from our previous transcriptome data; and the other purpose is that using the newly designed SSR markers to distinguish the duplications or synonymy garlics and to evaluate the genetic diversity of a garlic germplasm collection including 127 accessions.

2. Materials and methods

2.1. Plant materials and morphological data collection

A total of 127 garlic accessions kept and propagated by the Vegetable Physiology and Biotechnology Laboratory, Northwest A&F University, were used for genetic diversity and population

structure analysis. The detailed names and geographical origins of these garlics were listed in Table S1. These accessions were field-grown at the Wuquan Experiment Station (WES-NWAFU) of Northwest A&F University (108°08' E, 34°29' N) in Yangling, Shaanxi Province, China. These garlic accessions were grown in two growing seasons of 2019 and 2020, which were both sown in the September and harvested in April or May of the following year. About 500 individuals were planted for each garlic accession and the standard horticultural practices were performed following Liu et al. (2019). Thirty typical individuals were randomly selected for morphological data collection. The growth period of each garlic accession was recorded as the days from sowing to the harvest of mature bulbs. Garlic bulb related morphological traits, including bulb height (the height from the base of a bulb to the highest point), bulb width (maximum transverse width of a bulb), bulb weight (the average weight of 30 typical bulbs for each accession), rind color, and the clove number per bulb, were also recorded. All the collected phenotypic data were also listed in Table A.

2.2. Identification of SSR loci and marker development

We previously assembled and annotated 289,142 unigenes using a garlic transcriptome data set (Liu et al., 2020). Here, the simple sequence repeats or microsatellites (SSRs) were identified among these garlic unigenes using a computer program MISA (MIcroSATellite identification tool) version 1.0. The default parameters were used to screen a minimum of 6 repeats for dinucleotide motifs and 5 repeats for trinucleotide, tetranucleotide, pentanucleotide, and hexanucleotide motifs. For those detected putative SSRs, the Primer 3.0 software was used to design flanking primers with the corresponding unigene sequences following the criteria used in Liu et al. (2015). To assess the qualities of these newly developed SSR primer pairs, fifty SSR primer pairs were randomly selected to evaluate the genetic diversity and population structure of a panel of garlic germplasms that includes 127 accessions collected worldwide. Besides, ten polymorphic garlic SSR markers previously used in our laboratory and seven SSR markers used by Ma et al. (2009) were also selected to genotype these garlic germplasms. The detailed sequence information of all SSR markers used and developed in this study are listed in Table 1 and Table S2.

2.3. DNA extraction, PCR amplification and genotyping

About 0.25 g freeze-dried young leaf samples were weighted and ground into a fine powder for each garlic accession. The total genomic DNA of each sample was extracted using the CTAB procedure (Murray and Thompson, 1980). The quality and quantity of extracted DNAs were examined by electrophoresis in 1% agarose gel and measured using NanoDrop 2000 spectrophotometers, respectively. The high-quality DNAs were diluted into 50 ng/μl with ddH₂O and stored at -20 °C until further use.

PCR amplifications were carried in 10 μl reactions, each containing 2 μl template DNA (50 ng/μl), 0.5 μl forward primers (5 μmol/l), 0.5 μl reverse primers (5 μmol/l), 2 μl ddH₂O and 5 μl 2 × Taq PCR Master mix (Tiangen Biotech Co., LTD, Beijing, China). The PCR program was as follows: 5 min at 95 °C, 40 s denaturing at 94 °C, 40 s annealing at 68 °C and 40 s elongation at 72 °C, followed by a 2 °C reduction in the annealing temperature per cycle for 6 cycles. Then reduce annealing temperature in each cycle by 1 °C for 8 cycles from 58 °C; the annealing temperature was maintained at 50 °C for the remaining 20 cycles, followed by a final step at 72 °C for 5 min. The amplified PCR products were separated by vertical electrophoresis on 8% polyacrylamide gel in 1 × TBE buffer at a

Table 1
Characteristics and diversity statistics for 29 SSR loci evaluated in 127 garlic accessions.

Locus name	Primer sequence 5'- 3'	Chromosome	Physical position (bp)	<i>Na</i>	<i>He</i>	<i>H</i>	PIC
XNGSSR0001 ^a	F: ACAACACTATGATTGATGGGGT R: ATGGGGCAATCCAATGGTGT	8	226328573	3	0.94	0.01	0.44
XNGSSR2730 ^a	F: GTGCCAATCATGATACCACG R: GCACCTCACCTTCTCGAA	6	585270159	2	0.22	0.50	0.28
XNGSSR2855 ^a	F: AGAACCTGCCACATACGACG R: CTTTGAACCGGACATGCTGC	5	1909747304	2	0.83	0.16	0.38
XNGSSR3540 ^a	F: GGCCTGATCGTATTCAAAGTGG R: TCGATGAGGTATCAAGAACATGGA	5	701048647	2	0.91	0.10	0.39
XNGSSR1498 ^a	F: GCTAAATTCAGAGCACCGCC R: ACGTCCGACTGAGAACGTTT	1	166628208	2	0.94	0.19	0.38
XNGSSR 1845 ^a	F: CGACGAGGAAAACACGAGAGA R: TGCAGTGATGTTTATGCGTGT	4	696164998	4	0.13	0.23	0.25
XNGSSR4024 ^a	F: TGCAAATCAGGATCTGCCTT R: AATATTTGGCCTTTTGCCCC	6	456640066	3	0.06	0.10	0.24
XNGSSR4161 ^a	F: GGCTGCGGTGAGTAGGAAA R: TGCATCTCGAATCTGAAATCTAGAGT	3	1765423781	6	0.03	0.46	0.22
XNGSSR0546 ^a	F: ACCAGTGATGAGCTGCGATT R: CAGAAGGAAAACATTCGGGCA	3	519426765	2	0.92	0.08	0.37
XNGSSR1236 ^a	F: GAGAGGGAGGGAGGGGGA R: TGTGACTGGCCTTGCCATG	1	206936703	3	0.46	0.28	0.35
XNGSSR1270 ^a	F: TGTCTCGTGGATTCCTGCTG R: CACAACCAATGCCCAGCAG	8	367723659	2	0.94	0.33	0.39
XNGSSR1612 ^a	F: GCGAACAAGGTTTGGAGCGT R: ATGGAGCAACGATGACAGCA	4	833765837	6	0.88	0.32	0.39
ES-025 ^b	F: TTGCTCTCAATGCCTCAAA R: TGCTACTACCAAGTATGGCTTCA	5	119688378	2	0.92	0.08	0.37
ES-027 ^b	F: AACAGGTGAACAGGGAGACA R: GGCATATTTGATGGAGATAGGG	7	1477657420	3	0.93	0.17	0.45
ES-040 ^b	F: TTTCCACCGTTAAATTTGTGC R: ATGGGAAACACCGTGAACAG	8	1294124402	2	0.14	0.47	0.22
ES-041 ^b	F: TGCCGTAATGGAATTCCTG R: TGATAGGTGGAGAGTACAGTCGTG	1	766185733	5	0.64	0.21	0.36
ES-043 ^b	F: CATGCTGTAGAGTTTGGCATTG R: GTCGCCACAAAACAAAGTGT	5	65932939	4	0.83	0.37	0.39
ES-044 ^b	F: TTTCCCTTTCGTTTGAAT R: AACCTTGTGATAATTGAAAAATATGC	4	975639591	4	0.54	0.37	0.33
ES-045 ^b	F: AAAGTTGGGAGGGAATTACA R: GCAAGTCAATCAGTCGATGTG	7	1381636957	3	0.91	0.22	0.38
ES-046 ^b	F: TGATTTTGATAATTTGTCTCCCTT R: CCCACAACCGAAATAAGCA	7	79317313	4	0.43	0.35	0.45
ES-085 ^b	F: GGGGAGGAGAGGAGGTTAGT R: TTTGTGTTGCATGTGTGGAA	5	283464326	5	0.95	0.32	0.41
ES-087 ^b	F: TACCTGTTGGGAGCAGATTTC R: GTCACCGAGTCAGTGGGAAG	1	644066678	4	0.78	0.24	0.44
GB-ASM-040 ^c	F: CACAGCAACATGCACCAT R: TGCCGGAACCTCGATATT	3	883255742	4	0.72	0.21	0.41
GB-ASM-053 ^c	F: ACAAGGTCGACATCGTTTG R: GGGCTTACCTGAACACA	1	188253020	5	0.78	0.25	0.46
GB-ASM-059 ^c	F: CTTGCCGGAACCTCGATATT R: CACAGCAACATGCACCAT	1	14861506	4	0.54	0.36	0.27
GB-ASM-072 ^c	F: CACGGAATCTTCTTGG R: TGCAAAGCAATATGGCAG	7	603543167	6	0.81	0.23	0.37
GB-ASM-078 ^c	F: TGTTCCAACAGATTTAATGC R: AAGTGGCGTTGTGCTG	7	182069094	3	0.8	0.27	0.49
GB-ASM-080 ^c	F: AATCTCCCTCAAAGTCCC R: CCTGTATTTGTGTAAGCATCA	1	423593679	2	0.85	0.25	0.38
GB-ASM-109 ^c	F: GGTCTCTCATCCACCGT R: GTGTGGGCATGATTGAC	7	844712803	4	0.72	0.24	0.39
Mean				3.48	0.67	0.25	0.37

Note: Letters 'F' and 'R' precede the nucleotide sequence of forward and reverse primers, respectively. *Na*: number of alleles per loci; *He*: expected heterozygosity; *H*: Nei's gene diversity (Nei, 1973); PIC: polymorphism information content.

^a New garlic SSR markers developed in this study.

^b Previously developed SSR markers by our research team.

^c SSR markers developed by Ma et al. (2009).

constant 180 V for 1 h, visualized with silver staining, and photographed with a digital camera. The clear and unambiguous bands of all the polymorphic SSR markers were scored for 127 garlic accessions and calculated into co-dominant genotypic matrix in GeneAlix6.5 (Peakall and Smouse, 2012) following Zhu et al. (2016), which was used for the following data analysis.

2.4. Data analysis

The ANOVA of phenotypic traits between the years 2019 and 2020 was conducted first with SPSS 21, and then the 'pheatmap' R package was used to conduct a phenotypic cluster analysis of the 127 garlic accessions. In addition, the 'pairs. Panels' function in

'psych' R package (Northwestern University, Evanston, IL, USA) was used for analyzing the correlations among phenotypic traits (<https://cran.r-project.org/web/packages/psych/citation.html>).

General statistical parameters that reflect the genetic diversity of the 127 garlic accessions, according to the summarization of Pagnotta (2018), the number of alleles (N_a), expected heterozygosity (H_e), and Nei's gene diversity (H) were determined using the software of Popgene (v.1.32). The polymorphism information content (PIC) was calculated with PowerMarker (v.3.25). The pairwise genetic distances were used to examine the hierarchical clustering of garlic accessions via a dendrogram based on the unweighted pair-group method with an arithmetic mean (UPGMA) analyses and the neighbor joining tree was constructed by NTSYS 2.10 software.

The GenAIEx 6.5 was used to estimate the analysis of molecular variance (AMOVA). The redundant garlic accessions were identified by genetic distance ($GD = 0$), and the GD was calculated with NTSYS 2.10 software. Further analysis was performed using 102 garlic accessions after removing the redundancy.

For population structure analysis, the Bayesian model-based clustering was performed in STRUCTURE v.2.3.4 (Pritchard et al., 2000) to infer the appropriate clusters (K), using a burn-in of 10,000, run length of 150,000, and assuming admixture model and correlated allele frequencies. Ten runs of STRUCTURE were performed by setting the number of populations (K) from 1 to 10. The most probable value of K was select by ΔK methodology using the web-based software STRUCTURE HARVESTER v.0.6.94 (Earl and VonHoldt, 2012).

The correlation between the cophenetic matrix of Euclidean distances (morphological traits) and the cophenetic matrix of genetic distances was also calculated with GenAIEx 6.5 through the Mantel test. The comparison of morphological traits among different cluster accessions was conducted with ANOVA used SPSS 21, in which the P -value = 0.05 was used as a cut-off for the significant difference.

3. Results

3.1. The diversity of 127 garlic accessions revealed by the investigated morphological traits

Morphological traits, including growth period, cloves number, bulb height, width, weight, and rind color of 127 garlic accessions, were listed in Table S1. The ANOVA showed no significant differences for these observed phenotypic traits between the two growing seasons of 2019 and 2020 (Table S3).

Based on the observed data of 2020, this garlic germplasm collection had a relatively higher diversity among the investigated traits (Table S1, and Fig. 1). For the rind color of garlic bulbs, eighty-two accessions had purple rinds, while all the rest showed white rinds. Among the 127 garlic accessions, bulb height ranged from 24.4 mm (GS018 and GS112) to 45.0 mm (GS025), with an average height of 33.3 mm. Both GS018 and GS025 showed the narrowest (28.4 mm) and widest (62.7 mm) bulb width, respectively, with the mean bulb width of 43.8 mm. The minimum clove number was 4.7 for GS088, while the highest was 16.0 for GS024, with an average of 11. The bulb weight exhibited the largest variation ranging from 7.2 g for GS023 and GS111 to 70.1 g for GS024. The earliest-mature accession was GS044 with 246 days growth period, while the latest-harvest garlics were at 290 days for GS011, GS012, GS013, GS014, GS016, GS035, GS083, and GS118. In addition, for each accession, the bulb width was larger than the bulb length, which suggested that all these garlic accessions are characterized with flat shape bulbs.

Except for the rind color, the distributions and Pearson correlations of these investigated traits were calculated and displayed in Fig. 2. As shown in Fig. 2, the growth period, clove number, bulb height, width, and weight were all continuously distributed over the 127 accessions. The distributions of most accessions were within the ranges of 32–38 mm for bulb height, 35–50 mm for bulb width, 10–14 for clove number, 20–40 g for bulb weight, and 260–270 days for growth period. According to the calculated correlation coefficients, bulb weight had significant positive correlations with bulb height, width, and clove number, especially for the bulb width, which had a very strong correlation. Besides, clove number, bulb height, and width showed significant positive correlations with each other. However, growth period only had a weak correlation with bulb height.

Clustering analysis of these 127 garlic accessions was performed based on the investigated morphological traits. The phenotypic-based dendrogram (Fig. 3) illustrates that all the garlic accessions were classified into three major groups (Cluster I, II, and III). The majority (116 accessions) of this garlic collection were grouped into Cluster III, with only three accessions included in Cluster I, of which two (GS024 and GS025) were from Belgium, and one (GS120) is from the Xinjiang Province of China, and the other 8 accessions were grouped as Cluster II, which from America (GS018), Japan (GS022), Thailand (GS023) and China (GS077, GS088, GS108, GS111, GS112).

3.2. Distribution of SSR motifs and primer pairs development

A total of 4372 SSR loci were detected from assembled sequences of 289,142 unigenes. The frequencies, types, and distributions of these SSRs were analyzed and shown in Fig. 4. For these 4372 SSR loci, the dinucleotide repeat motifs were the most abundant type (2,461, representing 56.3%), followed by trinucleotide repeat motifs (1,811, representing 41.4%) and tetranucleotide repeat motifs (91, representing 2.1%), while the distributions of pent nucleotide and hexanucleotide repeat motifs were relatively rare (9, representing 0.2%). In addition, the repeat times of different SSR motifs were mainly scattered from five to ten. Among these identified SSR loci, 83 motifs were detected, in which nearly half of these SSRs were derived from five types of dinucleotide repeat motifs, including TA/TA (893, representing 20.4%), TG/CA (551, representing 12.6%), AC/GT (516, representing 11.8%), AG/CT (246, representing 5.6%), and TC/GA (227, representing 5.2%).

The designed primer pairs of these 4372 SSRs are listed in the Table S2. The putative amplified product sizes of these SSR markers ranged from 100 to 280 bp with a mean size of 210.2 bp. These primers had the following features: the sequence lengths ranging from 18 to 27 bp, the GC contents varied from 25.9 to 77.8%, and the T_m changed from 57.0 to 62.9 °C. The averaged sequence length, GC content, and T_m of these SSR primers were 20.6 bp, 50.8%, and 59.6 °C, respectively. Furthermore, compared with the garlic SSR markers developed by Liu et al. (2015), 244 consistent markers were found among the markers of this study (Table S2). This may suggest that most SSR markers of the present work are newly developed markers for garlic.

3.3. SSR marker assay and their informativeness

To assess the qualities of the newly developed SSR markers, 50 pairs of SSR primers were randomly selected to firstly genotype 10 of the 127 garlic accessions. Among these 50 SSR primer pairs, 39 SSR markers (78.0%) successfully amplified targeted DNA fragments, of which 12 SSR primer pairs exhibited polymorphism among these 10 selected garlic accessions. These 12 markers and

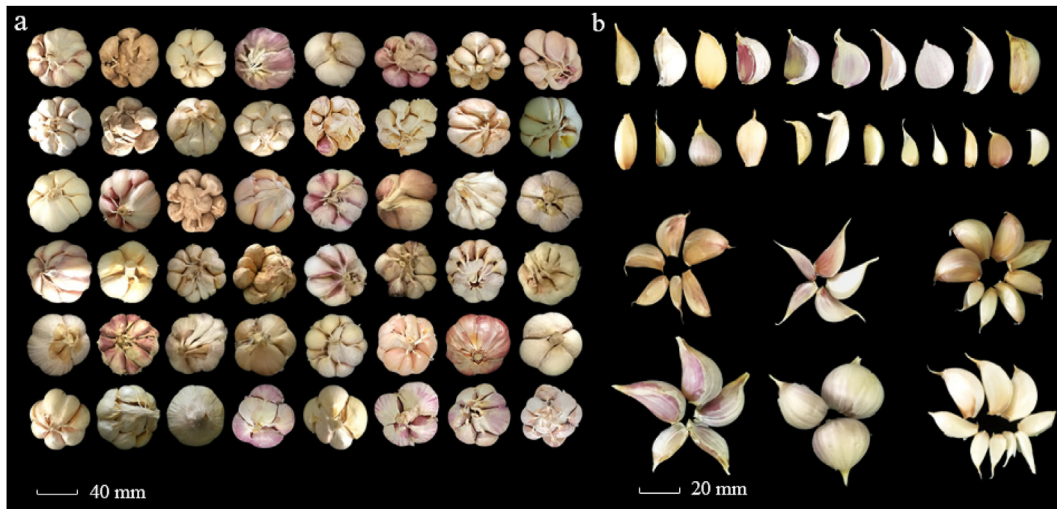


Fig. 1. Typical garlic bulbs (a) and cloves (b) with diverse phenotypes.

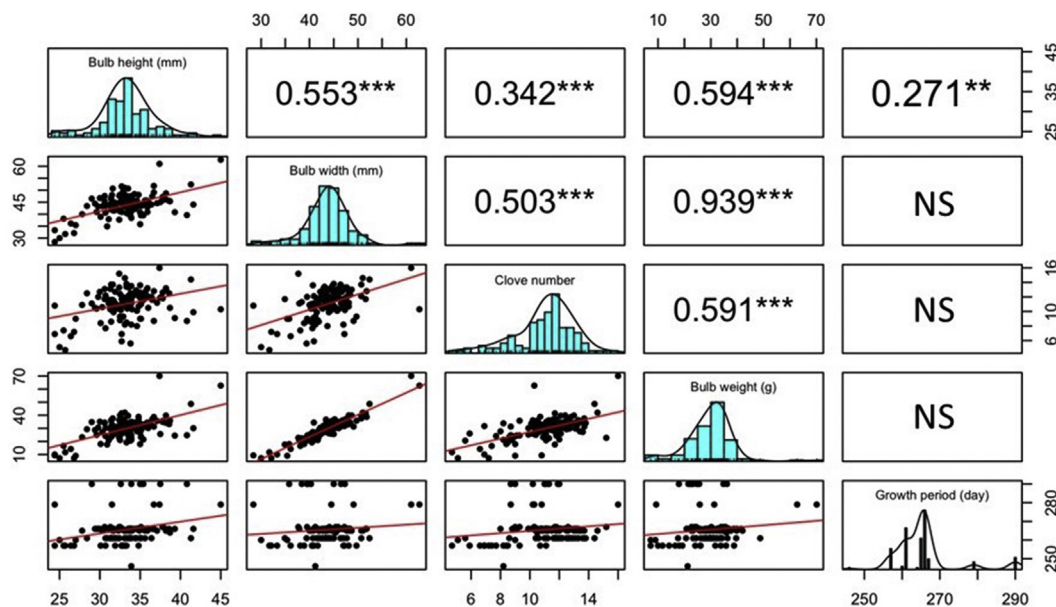


Fig. 2. Pearson correlation coefficients among phenotypic traits in 127 garlic accessions. NS, **, *** indicate not significant, significant at $p < 0.01$, $p < 0.001$, respectively. The bar charts show the value distribution of the phenotypic traits.

another 17 SSR markers listed in Table 1 were used together to genotype all the 127 garlic accessions. The genotyping analysis indicated a total of 79 alleles among these 127 accessions, with a range of 2–6 alleles and an average of 3.48 alleles per SSR marker (Table 1). Furthermore, the polymorphism information of SSR markers reflected by parameters expected heterozygosity (H_e), Nei's gene diversity (H), and the polymorphism information content (PIC) were also calculated and listed in Table 1. For these 29 markers, the H_e ranged from 0.01 to 0.95 with a mean value of 0.67. The H of each marker changed from 0.01 to 0.47, with the averaged $H = 0.25$. The PIC varied from 0.22 to 0.49, with an average of 0.36, which indicated that these SSR markers were good enough for evaluating the genetic diversity of the 127 garlic accessions.

3.4. The genetic diversity and population structure of the 127 garlic accessions

The genetic distance matrix of 127 garlic accessions was calculated based on the genotyping data of all 79 alleles. The neighbor-joining tree of these accessions was constructed with the obtained genetic distances by the UPGMA method, and the dendrogram was shown in Fig. 5. These 127 accessions were mainly grouped into three clusters, in which Cluster 1, 2, and 3 had 7, 15, and 105 garlic accessions, respectively. No direct relations were revealed between the geographical location of each of the garlic genotypes and their origin area. Although the garlic genotypes in Cluster 1 were all collected in China, they originated from entirely different locations of China, e.g. four accessions (GS081, GS083, GS113, and GS118)

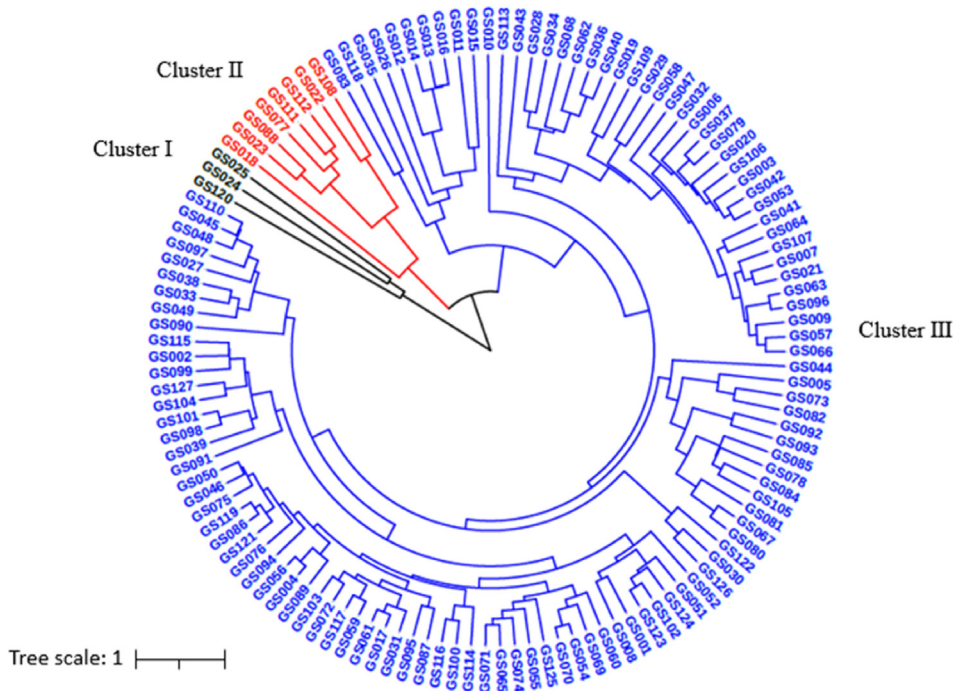


Fig. 3. Cluster dendrogram of 127 garlic germplasm based on phenotypic traits.

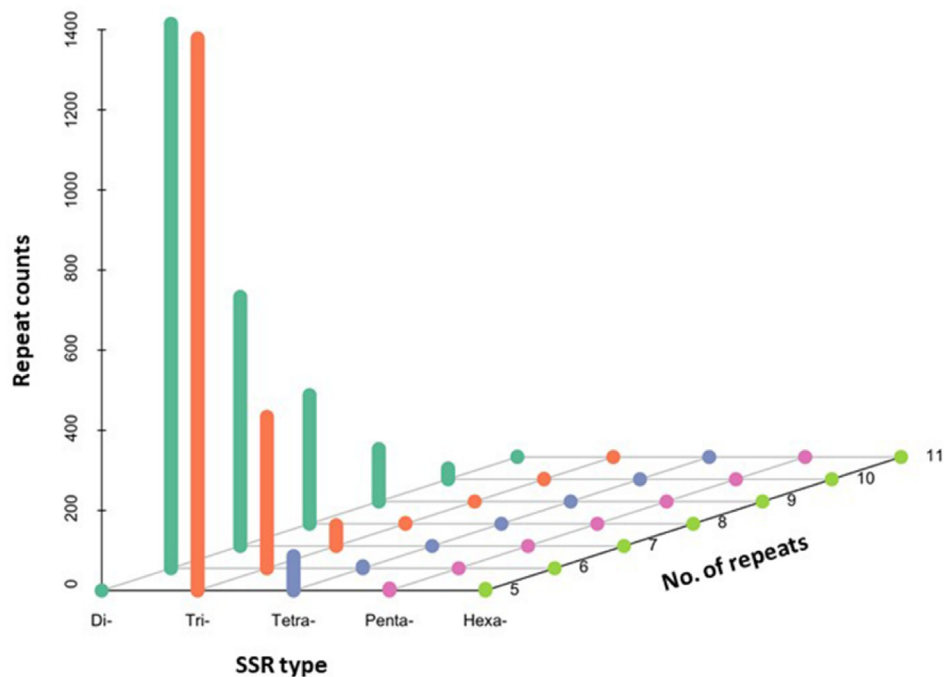


Fig. 4. Distribution of various SSR motifs with different numbers in transcriptome of garlic. The X-axis means nucleotide number in SSR repeat unit, Y-axis means the repeat number for a given unit length, Z-axis means the counts of specific repeat type.

were collected from the northwest of China, one (GS076) from the southwest of China, and the other two (GS087 and GS090) from the northeast and north of China, respectively. Similarly, the original locations of accessions in Cluster 2 were scattered at the east, southwest, central, north, and northwest parts of China, respectively. In Cluster 3, 78 garlic accessions were collected from different regions of China, and the other 27 genotypes were obtained from other 15 countries.

For further study of population structure and genetic variation, a total of 14 groups of potential redundancy garlic genotypes were identified from the 127 garlic accessions according to the genetic distance ($GD = 0$) and the SSR marker based dendrogram (Table 2). Group 4 had the highest number (8 garlic accessions) of redundancy genotypes, followed by Group 7, which had 4 garlic genotypes; three repetitive garlic genotypes were found for all of the Group 3, 8, and 11; while two redundancy garlic accessions per

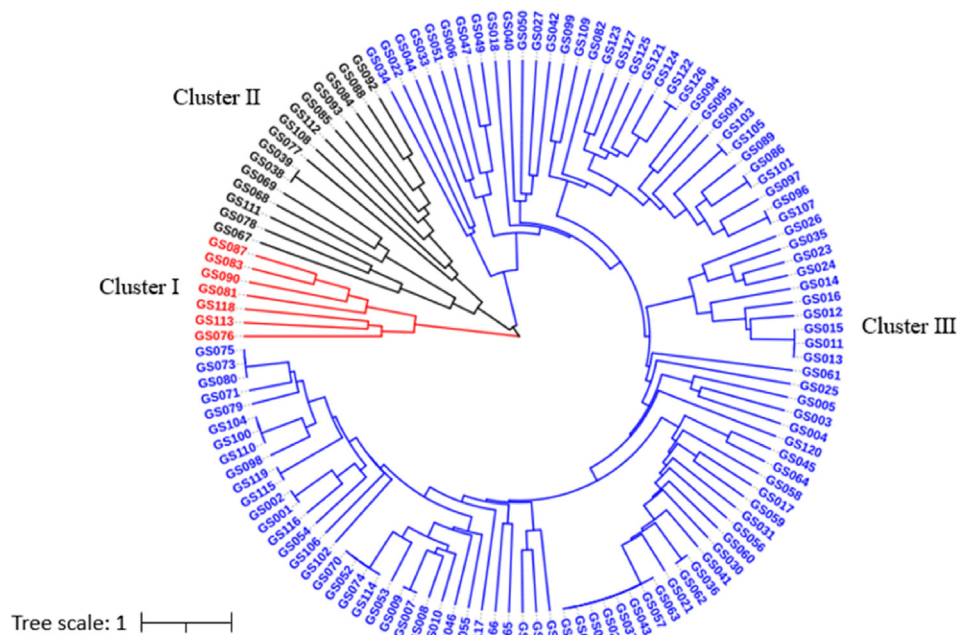


Fig. 5. UPGMA dendrogram demonstrating genetic relationships among 127 garlic accessions. The dendrogram is based on the Nei's genetic distance coefficient (GD).

Table 2
The redundant garlic accessions identified by the genetic distance.

Redundancy	Garlic accessions								
Group 1	GS001	GS002							
Group 2	GS007	GS009							
Group 3	GS011	GS013	GS015						
Group 4	GS019	GS020	GS028	GS029	GS037	GS043	GS057	GS063	
Group 5	GS036	GS062							
Group 6	GS038	GS039							
Group 7	GS052	GS070	GS074	GS114					
Group 8	GS073	GS075	GS080						
Group 9	GS086	GS101							
Group 10	GS096	GS107							
Group 11	GS100	GS104	GS110						
Group 12	GS103	GS105							
Group 13	GS115	GS119							
Group 14	GS122	GS126							

Note: The identify standard was the genetic distance (GD) = 0. The redundant garlic accessions showed in the same row.

group were identified for all the rest 9 Groups. Although some redundancy of garlic accessions had the same geographical origins, the redundancy genotypes originated from different provinces of China or from different countries were also common, such as the repetitive genotypes of GS001 vs GS002, GS007 vs GS009, GS011 vs GS013 vs GS015, and so forth. Overall, after eliminating these redundancy genotypes (25 accessions), 102 unique garlic germplasm were determined from the raw 127 garlic accessions.

After removing the 25 redundant accessions, the genetic population structure of the unique 102 garlic genotypes was further analyzed based on their genotypic data derived from 29 SSR markers. Analyses of K-value (number of clusters) from 1 to 10 revealed the highest peak of ΔK at K = 3, where there was no tendency to divide into subgroups (Fig. S1), which indicated that three clusters were involved in this garlic collection of 102 accessions. The result of genetic structure clearly showed three clusters, including Cluster 1, 2, and 3. For each cluster, most of the belonged accessions were from a single primitive ancestor with a few mixed individuals (Fig. 6). Furthermore, the structure distribution of three clusters of these 102 garlic accessions was highly consistent with

the cluster analysis of the whole collection of 127 garlic, in which the Cluster 1, 2, and 3 of the genetic structure analyses corresponded to the Cluster I, II, and III of the UPGMA-based dendrogram, respectively.

The population differentiation was evaluated through the analysis of molecular variance (AMOVA) utilizing the obtained data of SSR markers, in which the genetic variation of within-population was 90%, and that of among populations was only 10% (Table 3). In addition, the overall *Fst* value (0.18) indicated a high level of genetic differentiation among the collection of 102 garlic accessions, according to Wright (1978) who defined the genetic differentiation as low for *Fst* < 0.05, moderate for 0.05 < *Fst* < 0.15, high for 0.15 < *Fst* < 0.25, and very high for *Fst* > 0.25.

3.5. The relationship between the genetic differentiation and morphological traits of garlics

The correlations between genetic distance and phenotypic traits were analyzed by the Mantel test (Table 4). The genetic distance had weak correlations with all the investigated morphological

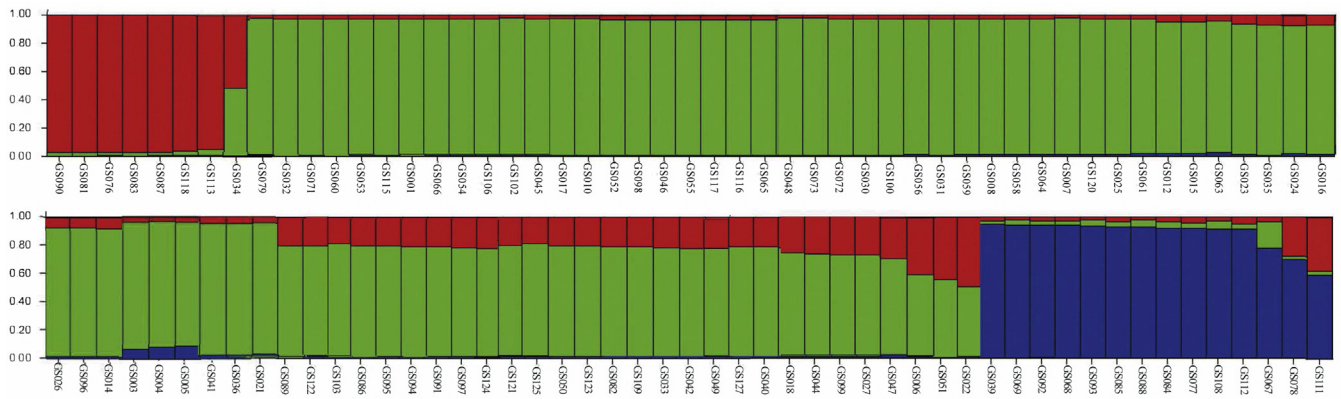


Fig. 6. Genetic structure of 102 garlic accessions as inferred by STRUCTURE based on 29 SSRs. Each accession is represented by a bar. The number below each vertical bar corresponds to the number of each accession ID in the Table S1. Identified subgroups are cluster 1 (red color), cluster 2 (green color), cluster 3 (blue color).

Table 3
Analysis of molecular variance (AMOVA) of 102 garlic accessions.

Source	Degrees of freedom (df)	Sum of squares (SS)	Mean squares (MS)	Estimated Variance	Percentage of variation	Stat	Value	P (rand ≥ data)
Among population	2	81.53	40.76	1.53	10%			
Within-population	99	1370.00	13.84	13.83	90%			
Total	101	1451.53	15.36	15.36	100%	Fst	0.18	0.001

traits, in which the growth period ($r = 0.13$) showed the lowest correlation with genetic distance, while the correlation between the clove number and the genetic distance was the highest ($r = 0.37$). The correlation coefficients of bulb weight, height, and width were 0.16, 0.21, and 0.15, respectively. In addition, the potential correlation between the genetic distance and geographical distance was also analyzed, while no significant correlation was found for these 102 garlic accessions.

The potential differences of morphological traits among the three clusters were further analyzed (Table 5). The results showed that all the traits of Cluster 3 were significantly higher than those of Cluster 2, except for clove number. The traits of Cluster 1 were also significantly higher than those of Cluster 2, when compared Cluster 1 and Cluster 3. Only the clove number showed a significant difference, where the accessions of Cluster 3 had more cloves per bulb. Overall, both the Cluster 1 and Cluster 3 were relatively superior to the Cluster 2 in phenotypic traits. Besides, when considering the rind colors of garlic bulbs, all these three clusters did not show any specificity since each cluster scattered with both white and purple rind colors.

4. Discussion

Although most garlic varieties propagate asexually, they exhibit considerable morphological differences within and between them (Bradley et al., 1996). Evaluating the genetic variation of cultivated garlic accessions is helpful for phenotypic identification and core germplasm construction. In addition, the analysis of genetic

diversity and relatedness between accessions are also important for garlic selection and breeding purposes (Figliuolo et al., 2001). In early studies, the variation of different garlic clones was mainly investigated in morphology and enzymology (Al-Zahim et al., 1997; Ipek et al., 2003). Highly morphological diversity was observed in garlic from several phenotypic traits, including the bulb size, shape and rind color, clove rind color, bolting habit, and clove number and weight (Singh et al., 2014; Raja et al., 2017). In the current study, the 127 garlic accessions exhibited relatively higher diversity among the investigated traits where the clustering analysis divided these accessions into three clusters. However, based on the morphological traits only, the potential redundant accessions were not accurately identified, which might indicate the shortage of genetic diversity evaluation of garlic germplasms from the aspect of morphology. Polyzos et al. (2019) reported that the garlic diversity examined based on morphological traits is much dependent on the genetic composition and the environmental conditions such as cultivation practices, soil properties, and fertilizing regimes. This might explain the major reasons for the disadvantage of morphological traits based genetic diversity analysis in garlic.

For genetic diversity evaluation, SSR markers have been widely applied in many plant species to evaluate genetic diversity, to construct genetic maps, and to determine species lineages. However, the insufficient number of SSR markers is a major obstacle for the related genetic studies of garlic. In the current study, 4372 SSR markers were newly developed from the sequences of 289,142 garlic unigenes that were assembled from our previous transcriptome data (Liu et al., 2020). Among these SSR markers, the

Table 4
The correlations between the genetic distance and geographical distance and phenotypic traits based on the Mantel test.

	GD vs Geographical distance	GD vs Bulb weight	GD vs Bulb height	GD vs Bulb width	GD vs Clove number	GD vs Growth period
Correlation value	0.02	0.16 ^b	0.21 ^b	0.15 ^a	0.37 ^b	0.13 ^b
p-value	0.07	0.00	0.00	0.02	0.00	0.00

Note: GD: genetic distance.

^a Significant correlations at $p \leq 0.05$ level.

^b Significant correlations at $p \leq 0.01$ level, p-value calculated using the distribution of correlation value estimated from 9999 permutations.

Table 5
ANOVA of phenotypic traits and growth period in different clusters.

Cluster	Bulb height (mm)	Bulb width (mm)	Clove number	Bulb weight (g)	Growth period (d)
1	36.4 ± 1.6 a	41.1 ± 1.2 a	9.7 ± 0.4 b	26.6 ± 2.1 a	272.4 ± 4.6 a
2	30.3 ± 1.0 b	38.5 ± 1.1 b	8.1 ± 0.6 b	19.9 ± 2.4 b	260.0 ± 1.1 b
3	33.6 ± 0.4 a	44.5 ± 0.5 a	11.5 ± 0.2 a	31.3 ± 1.0 a	266.0 ± 0.9 a

Note: Values are means ± standard deviations. Different letters indicate significant differences between classes of a given trait, at $p \leq 0.05$ (Tukey HSD test).

dinucleotide repeat motifs were the most abundant type, followed by trinucleotide repeat motifs and tetranucleotide repeat motif. This is different from the results of Liu et al. (2015), who reported that the trinucleotide repeat motifs were the most abundant type, followed by the dinucleotide repeat motifs. This might be mainly due to the SSR search criteria, the size of the dataset, and the database-mining tools (Varshney et al., 2005; Aggarwal et al., 2007). Besides, there were only 244 consistent SSR markers shared between Liu et al. (2015) and those reported here, which might also verify the potential differences of the assembled unigene sequences of these two studies. To check the qualities of these newly developed SSR markers, 50 pairs of SSR primers were randomly selected to genotype 10 garlic accessions, in which 78% markers (39) successfully amplified target bands, and the remaining 22% SSR markers (12) did not amplify any fragments. The failed amplifications of these SSRs may be probably due to the fact that the primers were designed across splice sites or large introns (Varshney et al., 2006; Cloutier et al., 2009; Liu et al., 2013).

Twelve polymorphic markers from those 39 newly developed and verified SSR markers were used to evaluate the genetic diversity of 127 garlic accessions together with another 10 EST-SSR markers previously developed by our lab and 7 SSR markers reported by Ma et al. (2009). The averaged PIC value of these 29 SSR markers was 0.36 with a range from 0.22 to 0.49, which is similar to the mean PIC value (0.38, ranging from 0.30 to 0.54) of 10 EST-SSRs used by Barboza et al. (2020) for assessing the genetic diversity of a collection of 73 garlic accessions. However, Ipek et al. (2015) obtained the mean PIC value of 0.60 for 26 EST-SSRs used in 31 garlic accessions. A higher mean PIC value in Ipek et al. (2015) probably caused by the 31 garlic were intentionally selected based on their previous AFLP analysis work (Ipek et al., 2003) to maximize the genetic variations, whereas no pre-selection and classification were performed for the 127 accessions used in the present research. In addition, the numbers or types of used markers and the population size or the actual genetic variation of an evaluated germplasm collection might all contribute to the changes of the PIC values (Barboza et al., 2020).

When assessing the genetic diversity and population structure of a germplasm collection, the SSR or other DNA markers were able to overcome the problem in the nomenclature of garlic accessions and the redundancy. The guaranteed identification of potential duplicates is helpful for reducing costs in the maintenance of garlic germplasms. Based on the genetic distance and the SSR marker-based dendrogram, 25 potential duplicates were identified among the 127 garlic accessions (Table 3). These redundancy garlic accessions may be due to the farmers in different areas tend to exchange their garlic varieties/produce from year to year, to protect yield against degeneration caused by continuous cropping using the same variety, and each farmer might name the same garlic variety with obviously different names; thus, some different garlic varieties were improperly given a same name (Wang et al., 2016). In addition, both the cluster analysis of 127 accessions and the population structure analysis of 102 unique garlic accessions revealed that three major clusters existed for the garlic germplasms used in this study. However, the garlic accessions that grouped into the

same cluster were not completely consistent with their geographical origins. Moreover, the Mantel test results also verified that there was no significant correlation between genetic distance and geographic distance, which was also found by Ipek et al. (2003), Pooler and Simon (1993), Wang et al. (2016), and Morales et al. (2013). Furthermore, many of the garlic accessions may have secondary source data rather than original wild collection data associated with them, which makes it difficult to trace their geographical origins (Volk et al., 2004). All the above issues might lead to the high redundancy accessions among the worldwide collections of garlic germplasms. In the present study, AMOVA analysis indicated 90% of the variation due to differences within-population variation, and only 10% of the variation was due to differences among populations, which might suggest the presence of genetic structures. This was consistent with the results of Zhao et al. (2011), who observed that 84.4% variation was from within-population differences and 15.6% variation was due to the divergences between populations. However, the results of Barboza et al. (2020) showed that the variation within-population and among populations was 72% and 28%, respectively. The proportional difference of variation sources might have been caused by the significantly different garlic accessions used in each of the related studies. In addition, for the 102 unique garlic accessions of this study, a higher level of genetic differentiation was indicated by the overall *Fst* value (0.18), which might suggest that these accessions are valuable for future breeding of new garlic cultivars.

Through the ages, garlic is asexually propagated by using their cloves since the lack of fertile and seed setting germplasms, which leads to that the introduction is the only effective way for obtaining relatively good varieties in the cultivation of garlic. The extensive and frequent introduction of garlic cultivars from different areas results with serious homonym and synonym issues for garlic accessions. Efficiently analysis of the genetic diversity and relatedness of different garlic cultivars, using the SSR or other molecular markers from the genomic DNA level, is of important for identification of potential redundant accessions and construction of core garlic germplasm panels, which could further reducing the maintenance cost of garlic germplasms and promoting the selection and breeding of high-quality garlic cultivars. Here, in this study, 25 duplicated cultivars were successfully identified from 127 garlic accessions based on the genotyping matrix of 29 SSR markers, which indicates the effectiveness and advantages of SSR markers in the identification of garlic germplasms. The relatively higher redundant garlic cultivars (19.68%, 25 of the 127 accessions) identified in this work also further confirmed the universal phenomena of homonym and synonym in garlic. Until now, the relatively large genome size (~16 Gb) and few available fertile germplasms are still restricting the processes of the molecular and genetic studies in garlic (Meryem et al., 2015). Although we newly developed more than 4000 SSR markers, given the large genome size of garlic, the available molecular DNA markers are still not enough. The newly released garlic genome might provide good opportunity for identifying and developing of new DNA markers from the genome-wide level. Then, thousands of new SSR or SNP markers will be available for using in the genetic studies of garlic. Anyway, at the present

stage, the newly developed SSR markers and the evaluated non-redundant garlic accessions are still valuable for promoting the related genetic studies in garlic and for the researchers, breeders, and producers to further utilize these garlic germplasms.

5. Conclusion

In conclusion, 4372 EST-SSRs were newly developed from assembled unigene sequences for garlic, in which 12 polymorphic markers combined with other 17 SSR markers were successfully used for the genetic diversity and population structure analysis of a garlic germplasm collection. There were 25 duplications among the evaluated 127 garlic accessions, and the rest of 102 unique garlic accessions can be divided into three clusters. Among this garlic germplasm collection, some interdependent traits such as bulb weight clove number, bulb height, and width exhibited significant positive correlations with each other. Overall, the newly developed SSR markers will be of great value for the related genetic studies in garlic, and the assessment of genetic diversity and population structure of the involved garlic accessions will also provide valuable information for further utilizing these garlic germplasms.

Author contributions

X. Li conducted the majority of the reported research. L. Qiao and B. Chen helped with genotyping of SSR markers. Y. Zheng, C. Zhi, and S. Zhang helped with the morphological data collection. Y. Pan and Z. Cheng conceived and supervised the research. X. Li and Y. Pan wrote the manuscript with the input of Z. Cheng. All authors reviewed and approved the final submission.

Declaration of competing interest

None.

Acknowledgments

This work was supported by the Education Development Fund of Northwest A&F University (2017) to Z. Cheng and the Chinese Universities Scientific Fund (2452019017) to Y. Pan.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.pld.2021.08.001>.

References

- Aggarwal, R.K., Hendre, P.S., Varshney, R.K., et al., 2007. Identification, characterization and utilization of EST-derived genic microsatellite markers for genome analyses of coffee and related species. *Theor. Appl. Genet.* 114 (2), 359–372.
- Al-Zahim, M., Newbury, H.J., Ford-Lloyd, B.V., 1997. Classification of genetic variation in garlic (*Allium sativum* L.) revealed by RAPD. *Hortscience* 32, 1102–1104.
- Avato, P., Miccolis, V., Tursi, F., 1998. Agronomic evaluation and essential oil content of garlic (*Allium sativum* L.) ecotypes grown in Southern Italy. *Adv. Hortic. Sci.* 12, 201–204.
- Baghalian, K., Ziai, S.A., Naghavi, M.R., et al., 2005. Evaluation of allicin content and botanical traits in Iranian garlic (*Allium sativum* L.) ecotypes. *Sci. Hortic.* 103, 155–166.
- Barboza, K., Salinas, M.C., Acuña, C.V., et al., 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. Hortic.* 261.
- Blair, M.W., Hurtado, N., 2013. EST-SSR markers from five sequenced cDNA libraries of common bean (*Phaseolus vulgaris* L.) comparing three bioinformatic algorithms. *Mol. Ecol. Resour.* 13, 688–695.
- Bradley, K., Rieger, M.A., Collins, G., 1996. Classification of Australian garlic cultivars by DNA fingerprinting. *Anim. Prod. Sci.* 36, 613–618.
- Cloutier, S., Niu, Z., Datla, R., et al., 2009. Development and analysis of EST-SSRs for flax (*Linum usitatissimum* L.). *Theor. Appl. Genet.* 119, 53–63.

- Cunha, C.P., Hoogerheide, E.S., Zucchi, M.I., et al., 2012. New microsatellite markers for garlic, *Allium sativum* (Alliaceae). *Am. J. Bot.* 99, e17–e19.
- Cunha, C.P., Resende, F.V., Zucchi, M.I., et al., 2014. SSR-based genetic diversity and structure of garlic accessions from Brazil. *Genetica* 142, 419–431.
- Earl, D.A., VonHoldt, B.M., 2012. Structure harvester: a website and program for visualizing structure output and implementing the Evanno method. *Conserv. Genet. Resour.* 4 (2), 359–361.
- Egea, L.A., Rosa, M.G., Andrzej, K., et al., 2017. Assessment of genetic diversity and structure of large garlic (*Allium sativum* L.) germplasm bank, by diversity arrays technology “genotyping-by-sequencing” platform (dartseq). *Front. Genet.* 8.
- Figliuolo, G., Candido, V., Logozzo, G., et al., 2001. Genetic evaluation of cultivated garlic germplasm (*Allium sativum* L. and *A. ampeloprasum* L.). *Euphytica* 121, 325–334.
- Govindaraj, M., Vetriventhan, M., Srinivasan, M., 2015. Importance of genetic diversity assessment in crop plants and its recent advances: an overview of its analytical perspectives. *Genet. Res. Int.* 2015, 431487–431487.
- Hong, C.J., Etoh, T., 1996. Fertile clones of garlic (*Allium sativum* L.) abundant around the Tien Shan Mountains. *Breed. Sci.* 46, 349–353.
- Ipek, M., Ipek, A., Almqvist, S.G., et al., 2005. Demonstration of linkage and development of the first low-density genetic map of garlic, based on AFLP markers. *Theor. Appl. Genet.* 110, 228–236.
- Ipek, M., Ipek, A., Simon, P.W., et al., 2003. Comparison of AFLPs, RAPD markers and isozymes for diversity assessment of garlic and detection of putative duplicates in germplasm collections. *J. Am. Soc. Hortic. Sci.* 128, 246–252.
- Ipek, M., Sahin, N., Ipek, A., et al., 2015. Development and validation of new SSR markers from expressed regions in the garlic genome. *Sci. Agric.* 72, 41–46.
- Kalia, R.K., Rai, M.K., Kalia, S., et al., 2011. Microsatellite markers: an overview of the recent progress in plants. *Euphytica* 177, 309–334.
- Lee, G., Kwon, S., Park, Y., et al., 2011. Cross-amplification of SSR markers developed from *Allium sativum* to other *Allium* species. *Sci. Hortic.* 128, 401–407.
- Liu, H., Deng, R., Huang, C., et al., 2019. Exogenous gibberellins alter morphology and nutritional traits of garlic (*Allium sativum* L.) bulb. *Sci. Hortic.* 246, 298–306.
- Liu, H., Wen, Y., Cui, M., et al., 2020. Histological, physiological and transcriptomic analysis reveal gibberellin-induced axillary meristem formation in garlic (*Allium Sativum* L.). *Plants* 9 (8).
- Liu, T., Zeng, L., Zhu, S., et al., 2015. Large-scale development of expressed sequence tag-derived simple sequence repeat markers by deep transcriptome sequencing in garlic (*Allium sativum* L.). *Mol. Breed.* 35, 204.
- Liu, T., Zhu, S., Fu, L., et al., 2013. Development and characterization of 1827 expressed sequence tag-derived simple sequence repeat markers in ramie (*Boehmeria nivea* L. Gaud.). *PLoS One* 8, e60346.
- Lu, J.J., Kang, J.Y., Feng, S.G., et al., 2013. Transferability of SSR markers derived from *Dendrobium nobile* expressed sequence tags (ESTs) and their utilization in *Dendrobium* phylogeny analysis. *Sci. Hortic.* 158, 8–15.
- Ma, K., Kwag, J., Zhao, W., et al., 2009. Isolation and characteristics of eight novel polymorphic microsatellite loci from the genome of garlic (*Allium sativum* L.). *Sci. Hortic.* 122, 355–361.
- Maaß, H.I., Klaas, M., 1995. Intraspecific differentiation of garlic (*Allium sativum* L.) by isozyme and RAPD markers. *Theor. Appl. Genet.* 91, 89–97.
- Meryem, I., Nihan, S., Ahmet, I., et al., 2015. Development and validation of new SSR markers from expressed regions in the garlic genome. *Sci. Agric.* 72, 41–46.
- Mohanty, P., Sahoo, L., Parida, K., et al., 2013. Development of polymorphic EST-SSR markers in *Macrobrachium rosenbergii* by data mining. *Conserv. Genet. Resour.* 5, 133–136.
- Morales, R.G.F., Resende, J.T.V., Resende, F.V., et al., 2013. Genetic divergence among Brazilian garlic cultivars based on morphological characters and AFLP markers. *Genet. Mol. Res.* 12, 270–281.
- Murray, M.G., Thompson, W.F., 1980. Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res.* 8, 4321–4325.
- Nei, M., 1973. Analysis of gene diversity in subdivided populations. *Proc. Natl. Acad. Sci. U. S. A.* 70, 3321–3323.
- Ovesná, J., Leisová-Svobodová, L., Kučera, L., 2014. Microsatellite analysis indicates the specific genetic basis of Czech bolting garlic. *Czech J. Genet. Plant Breed.* 50, 226–234.
- Pagnotta, M.A., 2018. Comparison among methods and statistical software packages to analyze germplasm genetic diversity by means of codominant markers. *J. I.* 1, 197–215.
- Peakall, R., Smouse, P.E., 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics* 28, 2537–2539.
- Polyzos, N., Papanotiropoulos, V., Lamari, F.N., et al., 2019. Phenotypic characterization and quality traits of Greek garlic (*Allium sativum* L.) germplasm cultivated at two different locations. *Genet. Resour. Crop Evol.* 66, 1671–1689.
- Pooler, M.R., Simon, P.W., 1993. Characterization and classification of isozyme and morphological variation in a diverse collection of garlic clones. *Euphytica* 68, 121–130.
- Pritchard, J.K., Stephens, M.J., Donnelly, P.J., 2000. Inference of population structure using multilocus genotype data. *Genetics* 155, 945–959.
- Rahman, K., Lowe, G.M., 2006. Garlic and cardiovascular disease: a critical review. *J. Nutr.* 136, 7365–7405.
- Raja, H., Ram, C.N., Sriombhargav, K.K., et al., 2017. Genetic variability assessment in garlic (*Allium sativum* L.) genotypes. *J. Pharmacogn. Phytochem.* 6, 1781–1786.

- Singh, L., Kaul, V., Gohil, R.N., 2014. Analysis of morphological variability in the Indian germplasm of *Allium sativum* L. *Plant Syst. Evol.* 300, 245–254.
- Vaek, J., ílová, Daniela, Melounová, Martina, et al., 2020. New EST-SSR markers for individual genotyping of opium poppy cultivars (*Papaver somniferum* L.). *Plants* 9, 10.
- Varshney, R.K., Graner, A., Sorrells, M.E., 2005. Genic microsatellite markers in plants: features and applications. *Trends Biotechnol.* 23, 48–55.
- Varshney, R.K., Grosse, I., Hahnel, U., et al., 2006. Genetic mapping and BAC assignment of EST-derived SSR markers shows nonuniform distribution of genes in the barley genome. *Theor. Appl. Genet.* 113, 239–250.
- Vavilov, N.I., 1951. The origin, variation, immunity and breeding of cultivated plants. *Chron. Bot.* 13.
- Volk, G.M., Henk, A.D., Richards, C.M., 2004. Genetic diversity among U.S. Garlic clones as detected using AFLP methods. *J. Am. Soc. Hortic. Sci.* 129, 559–569.
- Wang, H., Li, X., Liu, X., et al., 2016. Genetic diversity of garlic (*Allium sativum* L.) germplasm from China by fluorescent-based AFLP, SSR and InDel markers. *Plant Breed.* 135, 743–750.
- Wang, H., Li, X., Shen, D., et al., 2014. Diversity evaluation of morphological traits and allicin content in garlic (*Allium sativum* L.) from China. *Euphytica* 198, 243–254.
- Wright, S., 1978. *Evolution and the Genetics of Populations, Variability within and Among Natural Population.* The University of Chicago Press, Chicago, United States, p. 4.
- Zhao, W.G., Chung, J.W., Lee, G.A., et al., 2011. Molecular genetic diversity and population structure of a selected core set in garlic and its relatives using novel SSR markers. *Plant Breed.* 130, 46–54.
- Zhu, H.Y., Song, P.Y., Dal-Hoe, et al., 2016. Genome wide characterization of simple sequence repeats in watermelon genome and their application in comparative mapping and genetic diversity analysis. *BMC Genom.* 17, 557.