A Multi-marker Genomic Approach to Decipher the Divergence and Diversity in Selected Allium sativum L. cultivars

Narayana Chellaiya Johnson Packia Lekshmi
Noorul Islam Centre For Higher Education
Duraisamy Mahamuni
PSG College of Arts and Science
Johnson Raja Brindha
Noorul Islam Centre For Higher Education
Ramasamy Rajesh
rrajeshpbt@gmail.com

Pasumpon Thiru Muthuramalinga Thevar Memorial College https://orcid.org/0000-0002-4493-8995

Research Article

Keywords: garlic, ITS, barcoding, ISSR profiling, incomplete lineage sorting, breeding

Posted Date: March 11th, 2024

DOI: https://doi.org/10.21203/rs.3.rs-3978989/v1

Abstract

The genus *Allium* comprises plants of significant economic and medical importance, including onion, garlic, and leek plants. The genetic diversity of garlic plants (*Allium sativum*) is vital for improving agricultural practices, developing resilient crops, preserving genetic resources, and exploring the full range of culinary and medicinal potential within this important plant species. In this research, we investigated the results of genetic barcoding, focusing on the internal transcribed spacer (ITS) region; four distinct barcoding regions, matK, rbcL, and trnH-psbA; and the trnL and Inter Simple Sequence Repeats (ISSR) regions of *Allium sativum* L. (Amaryllidaceae), which were collected from three diverse cultivation sites. Our findings revealed significant interspecific diversity and intraspecific divergence among the three cultivars examined. Interestingly, the results from different genetic markers were consistent, with BDUT 1451 and 1452 consistently grouping together, while BDUT 1450 diverged. These findings emphasize the effectiveness of the multi-marker approach for exploring intricate genetic landscapes. Furthermore, they highlight the importance of genetic studies in understanding the diversity of breeding and the potential utility of this economically and medicinally important nutraceutical crop.

Introduction

Estimating Earth's species diversity has always been a contentious issue, with suggested numbers ranging from approximately 1.5 million (described thus far) to more than 5 billion species (Larsen et al. 2017). On the other hand, reports indicate that numerous species are disappearing and that more than a million are under the threat of extinction (Tollefson 2019). However, some studies suggest that nearly two-thirds of the existing species were described already (Appeltans et al. 2012; Costello et al. 2013; Li and Wiens 2023). Despite an increase in the number of taxonomists working on biodiversity, there has been a decline in the identification of new species (Costello et al. 2013). Notwithstanding considerable technological advancements in taxonomy, Engel et al. (2021) highlight persistent concerns regarding a dearth of qualified taxonomists, inadequate funding, insufficient training programs, and limited job opportunities within the field.

Conventional phytomorphological and plant anatomical studies limit our understanding of evolutionary taxonomy and species identification. These approaches involve four significant limitations. First, phenotypic plasticity and genetic diversity within the characteristics used for species identification can lead to erroneous determinations. Second, this method often overlooks phenotypically cryptic taxa, which are prevalent in many species groups (Knowlton 1993; Jarman and Elliott 2000; Aryakia et al. 2016). Third, phenotypic keys are often effective only for a specific life stage, leaving a substantial portion of individuals unidentified. Finally, while modern interactive tools represent a significant advancement, the use of keys often requires a high level of expertise, increasing the likelihood of misdiagnosis (Hebert et al. 2003). These limitations, along with the declining rate of new species identification, necessitate innovative approaches in taxonomic identification. In the late 1990s and early 21st century, the use of a microgenomic identification system emerged as a promising avenue for determining biological diversity. Leveraging the power of genetic analysis through the examination of

small fragments of an organism's genome (Hebert et al. 2003), this system utilizes genetic barcodes present in every cell, akin to universal product codes used for retail product identification. Despite the relative simplicity of determining life diversity through genomic barcoding, this approach provides a more efficient and precise means of categorizing and understanding the incredible variety of life forms on our planet.

DNA barcoding is a transformative methodology that empowers taxonomists by expanding their ability to complete a global inventory of diversity (Hebert and Gregory 2005). DNA barcoding not only benefits taxonomists but also scientists in diverse fields, such as forensic science, biotechnology, the food industry, and animal diet research (Creer et al. 2016; Deiner et al. 2017; Yang et al. 2020; Fernandes et al. 2021; Gostel and Kress 2022). In summary, genomic barcoding provides an expansive and versatile tool for untangling the intricate web of biological diversity, akin to how product barcodes streamline the identification of consumer goods. Efforts to develop universal genomic databases and accessions involve coordination among research organizations, government bodies, scientific societies, and others associated with genomic studies. Initiatives such as the Barcode of Life Initiative (BOLI), established in 2003, aim to improve the applicability and cost-effectiveness of genomic taxon identification (Costa and Carvalho 2010). Similarly, the International Barcode of Life (iBOL), established in 2008, seeks to transform biodiversity science by building DNA barcode reference libraries, sequencing facilities, informatics platforms, and analytical protocols through international collaboration (iBOL 2024).

Despite the potential benefits of DNA barcoding for both practitioners and users of classification, there is controversy within various systematic circles (DeSalle and Goldstein 2019; Honeycutt 2021; Pfeiler 2023). Although employing a single gene region as a DNA barcode does not guarantee complete taxonomic resolution, it offers a significant degree of proximity. Numerous studies on diverse animal groups suggest that DNA barcoding can achieve high levels of species resolution (Garg and Biju 2021; Dincă et al. 2023) and provides promising taxon identification and species determination in plant communities (Besse et al. 2021; Raskoti and Ale 2021; Sawarkar et al. 2021; Chac and Thinh 2023).

The Consortium of Barcode of Life (CBOL) recommends a 2-locus combination comprising the ribulose-1,5 bisphosphate carboxylase oxygenase large subunit (rbcL) and maturase K (matK) as the standard plant barcode (CBOL 2009). These two regions of chloroplast DNA were chosen for their efficient recovery of high-quality sequences and their ability to distinguish between species effectively (Burgess et al. 2011). In recent decades, researchers have increasingly utilized rbcL and matK sequences for both species identification (Starr et al. 2009; Asahina et al. 2010; Bieniek et al. 2015; Ho et al. 2021) and phylogenetic analysis (Asahina et al. 2010; Kuo et al. 2011; Maulidya et al. 2020; Cetiz et al. 2023). Additionally, the spacer between tRNA-His and photosystem II protein D1 (trnH-psbA spacer) and the ITS2 are widely employed for similar purposes (Chen et al. 2010; Gao et al. 2010; Fu et al. 2011; Bieniek et al. 2015; Intharuksa et al. 2020; Jiang et al. 2023). However, it is important to note that the application of universal barcode markers has its own limitations and is subject to further advances (Tekle and Wood 2018; Piper et al. 2019; Manzoor et al. 2022).

Accurate identification of crucial agricultural crops is beneficial for DNA barcoding, and insights into the evolutionary background of these crops have been obtained. However, the study of genetic diversity in *Allium* species has faced challenges due to the absence of readily portable codominant molecular markers during the early stages of this technology (McCallum 2007). While a variety of molecular marker methods have successfully addressed questions related to genetic diversity and species relatedness within *Allium* species, the search for robust and informative markers specific to *Allium* species has proven to be considerably more demanding (Xie et al. 2020).

Despite being economically important crops, some Amaryllidaceae crops, such as bulb onions and garlic, remain understudied compared to other major crops (Zhang et al. 2023). *Allium sativum*, an essential aromatic and nutraceutical ingredient, offers a wide array of health benefits (Wang et al. 2014; Rauf et al. 2022). Understanding the taxonomic evolution of functional traits requires exploring the molecular breeding and genetic diversity of this important food crop. With these considerations in mind, the present research aimed to explore the relationships among three distinct cultivars of garlic plants (*Allium sativum*) using ITS sequence and DNA barcoding, employing primers for matK, rbcL, trnH-psbA, and trnL. This study provides valuable insights into the genetic diversity and relatedness of these cultivars, contributing to our understanding of this important crop species.

Materials and Methods

We collected three distinct local cultivars of *Allium sativum* from different locations in Tamil Nadu, India: Poomparai (BDUT 1450) in Kodaikanal District, Vadugapatti (BDUT 1451) in Theni District, and Pannaikadu (BDUT 1452) in Kodaikanal District. Applying the CTAB method (Doyle and Doyle 1987), DNA extraction was conducted with previously washed and sterilized fresh, young roots, each weighing 200 mg per cultivar. Subsequently, a spin column kit was used to purify the DNA extracts from the garlic cultivars. Evaluation of the ethidium bromide-stained DNA extracts revealed the formation of bands upon UV exposure. To evaluate the sequence information of the chosen cultivars, we targeted specific regions: a spacer DNA (ITS) region, four candidate DNA coding regions (rbcL, matK, PsbA, and trnL) and an inter simple sequence repeater (ISSR) of the microsatellite region. Amplification of these regions was achieved through polymerase chain reaction (PCR) using the universal primers outlined in Table 1.

Table 1
Primers used in the barcoding and sequencing of *Allium* species

Genomic Target	Primers	Primer Sequence 5'-3'	Reference
Internal Transcribed Spacer Region	ITS1	TCCGTAGGTGAACCTGCGG	White et al., 1990
	ITS4	TCCTCCGCTTATTGATATGC	
Barcoding Regions	rbcLa-F	ATGTCACCACAAACAGAGACTAAAGC	Bafeel et al., 2011
	rbcLa-R	GTAAAATCAAGTCCACCRCG	
	MatK-3F KIM	CGTACAGTACTTTTGTGTTTACGAG	Costion et al., 2011
	MatK-1R KIM	ACCCAGTCCATCTGGAAATCTTGGTTC	
	psbA3_f	GTTATGCATGAACGTAATGCTC	Kress et al., 2005
	trnHf_05	CGCGCATGGTGGATTCACAATCC	
	trnL F	CGAAATCGGTAGACGCTACG	Taberlet et al., 1991
	trnL R	GGGGATAGAGGGACTTGAAC	
Inter Simple Sequence Repeats	ISSR8US	G(AACA)4	Jabbes et al. (2011)
	ISSR9	HVH(TC)8	
	ISHY 1b	(GA)8YT	
	ISHY 2	(AC)8YG	
	ISHY 3	(AG)8YT	
	ISHY 4	(CA)8RG	
	ISSR a	GCV(TG)8	
	ISSR8US	G(AACA)4	

For each PCR mixture, we included 10 µl of Taq premix, 4 µl of water, 1.5 µl of each of the forward and reverse primers, and 3 µl of DNA. The PCR Thermal Cycler Program was executed using an *Eppendorf ProS* (Hamburg, Germany). The thermal cycling programme for all sequence analyses comprised an initial cycle of 5 min at 94°C, followed by 35 cycles of 30 sec at 94°C, 30 sec at 58°C, and 60 sec at 72°C. The amplification was completed with a final cycle of 10 min at 72°C. To visualize the PCR products, all the samples were separated on a 1.2% agarose gel in 1× TAE buffer through electrophoresis at 60 V for 90 min. Subsequently, the separation was examined using a UV transilluminator.

Before sequencing, the PCR products were subjected to an additional purification step. The Sanger dideoxy method of DNA sequencing was used for the PCR products, and the obtained data were

imported and aligned using Molecular Evolutionary Genetics Analysis (MEGA v6.2.2). Key sequence statistics, covering nucleotide frequencies, transition/transversion bias (R = s/v), and variability across various sequence regions, were analysed. For sequence data analysis, both phenetic and cladistic methods were employed. The phenetic analysis utilized the neighbour–joining (NJ) method, while the cladistic approach employed the maximum parsimony (MP) method. These two complementary methods provided comprehensive insights into the genetic relationships and patterns present in the sequence data.

For each primer, the genetic sequence of the ISSR region was indexed to identify and compare polymorphic loci (P) among the cultivars. This information was subsequently utilized to compute the Shannon information index (I) (Lewontin 1972) and Nei's standard genetic distance (D) (Nei 1972). The allele diversity data obtained from these indices were used to evaluate the discriminative performance of each primer. Subsequently, the genetic distance information was used to construct a dendrogram with POPGENE v32 software. Finally, a population model was established through Bayesian analysis of the ISSR data using STRUCTURE v3.2.2 software.

Results and Discussion

Various primers were used to amplify and sequence distinct genomic regions of three chosen *Allium* cultivars. The number of base pairs in each sequenced region slightly differed among the three cultivars, with the trnL region exhibiting shorter lengths than the other barcoding regions analysed in this study (Table 2).

Table 2
Lengths of the sequenced genome regions of three selected *Allium sativum* cultivars for different molecular markers used in this study (number in base pairs).

Garlic Cultivar	Internal Transcribed Spacer Region	Barcoding Regions			
	ITS	matK	trnH-psbA	rbcL	trnL
BDUT 1450	645	851	650	537	287
BDUT 1451	651	853	646	540	291
BDUT 1452	630	851	640	547	259

The substitution pattern and rates for the ITS sequence were estimated using the Tamura and Nei (1993) model with a discrete gamma distribution to account for evolutionary rate differences among sites (5 categories, [+ G]). In the ITS region analysis, the shape parameter for the discrete Gamma distribution was estimated to be 200.0000. The mean evolutionary rates in these categories were 0.90, 0.96, 1.00, 1.04, and 1.10 substitutions per site, suggesting moderate evolution of these cultivars. The maximum log likelihood for this computation was − 2594.439, indicating a good fit with the Tamura–Nei model, significant heterogeneity, and a moderate evolutionary rate. The nucleotide frequencies of the ITS region observed in this Tamura–Nei model were A = 22.54%, T = 30.58%, C = 20.37%, and G = 26.51%. The

transition/transformation bias (R) for the ITS region was 1.09, indicating a bias higher than the expected level. Additionally, when the Kimura (1980) 2-parameter model was used for substitution patterns and rates, the nucleotide frequencies were A = 25.00%, T = 25.00%, C = 25.00%, and G = 25.00%. For estimating maximum likelihood (ML) values, a tree topology was automatically computed, and the maximum log likelihood for this Kimura 2-parameter computation was − 2617.110.

Both models exhibit a typical ITS region with AT richness, which might contribute to the higher R value than expected. When the phenetic neighbor–joining (NJ) method was applied to the ITS sequence, the phylogeny of the three cultivars resulted in a divergence of BDUT 1451 from that of the other two cultivars, BDUT 1452 and BDUT 1450, which were found to be in a separate cluster. In contrast, according to the cladistic MP method, BDUT 1450 and BDUT 1452 formed a single cluster, while BDUT 1451 diverged into a separate cluster (Figs. 1 and 2). The discrepancy in clustering patterns could conceivably be attributed to the high rate of heterogeneity, where the NJ method relies on overall sequence similarity, whereas the MP method focuses on shared derived characters. Interestingly, both methods consistently placed BDUT 1451 in a separate cluster, suggesting intraspecific divergence.

The phenetic neighbor-joining (NJ) method of ITS sequence analysis revealed that BDUT 1451 was closely linked to *A. sativum* gi228015332, and BDUT 1452 diverged from this cluster. BDUT 1450 was closely related to a cluster of *A. ampeloprasum* var. *harmone* gi338191565 and *A. ampeloprasum* var. *harmone* gi338191566. According to the cladistic method of ITS, BDUT 1450 and BDUT 1452 formed a single clade, whereas BDUT 1451 was closely linked to *A. sativum* gi228015332. The overall pairwise distance between them was 6.153. A dendrogram created by Baghalian et al. (2006) using 24 ecotypes revealed five main groups. Thirteen of the 24 ecotypes were included in cluster group A, two ecotypes in group B and seven ecotypes in group C. They found that there was no relationship between genetic divergence and geographical origin in the post planting stage, as genotypes from the same origin entered into different clusters, and genotypes from different origins also entered the same cluster. In the study by Figliuolo et al. (2001), molecular analysis confirmed the differentiation of *A. sativum* and *A. ampeloprasum* in the sequence cluster. The *A. ampeloprasum* cluster diverged from *A. sativum* through the use of single-nucleotide sequence variation in both the ITS1 locus and the RAPD locus. Sequence analysis of ITS1 DNA revealed detailed variation within the *A. ampeloprasum* complex and monomorphism within *A. sativum* in the study by Figliuolo and Di Stefano (2007). Several other studies have revealed intricate phylogenetic relationships, and the whole-genome sequence of Allium was generated through ITS analysis (Ipek et al. 2008, Hirschegger et al. 2010 and Nguyen et al. 2008).

One striking characteristic of the ITS data is the strangely large intrageneric genetic distances within *Allium* species. Kimura distances greater than 40% were found in the Friesen et al. (2000) study and in the Dubouzet and Shinoda (1999) study; Kimura distances often characterize the most distant genera within subfamilies or even families (Baldwin et al. 1995; Blattner and Kadereit 1999; Hsiao et al. 1999; Noyes and Rieseberg 1999). Intrageneric distances in other plant families are mostly less than 10% (Baldwin et al. 1995). These findings make *Allium* species either extraordinarily rapidly evolving taxa or ancient in origin, and molecular evolution is not accompanied by the increase in comparable numbers of

taxonomic categories. Moreover, a phylogenetic plot showed that all three cultivars belonging to the same species fell into distinct clusters, signalling their genetic divergence. Fredotović et al. (2014) performed phylogenetic analyses of the ITS1-5.8S-ITS2 region of 35S rDNA and the non-transcribed spacer (NTS) region of 5S rDNA of *A. xcornutum* and its relatives to the section *A. cepa*.

After the ITS region was sequenced, the phylogenetic relationships of the barcoding gene sequences, such as matK, rbcL, trnH-psbA and trnL, were determined via cladistic methods. In the analysis of the matK region, the shape parameter for the discrete Gamma distribution was estimated to be 0.0500. The mean evolutionary rates in these categories were 0.00, 0.00, 0.00, 0.03, and 4.97 substitutions per site. The fifth category stands out as remarkably divergent compared with the others. These results imply the presence of a conserved matK region with a low rate of evolution and significant intraspecific divergence. The nucleotide frequencies observed were A = 32.28%, T = 37.92%, C = 15.24%, and G = 14.57%, which confirmed the characteristic AT-rich chloroplast DNA. The maximum log likelihood for this computation was − 1210.897. Additionally, the estimated transition/transformation bias (R) is 0.11, and the maximum log likelihood for this computation was − 1287.698.

According to the phenetic neighbor–joining (NJ) method of phylogenetic analysis for matK, all three cultivars formed a single cluster, with BDUT 1450 and 1451 closely linked. According to the cladistic MP method, BDUT 1450 and 1452 were found in a single cluster that diverged from BDUT 1451 (Figs. 3 and 4). According to the NJ method of *matK* analysis performed in this study, all three cultivars were found in a single clade, and the overall pairwise distance between them was 0.012. Both phenetic and cladistic methods concur that BDUT 1451 diverged from BDUT 1450 and 1452. The matK region had a high evolutionary rate (4.97), suggesting the potential accumulation of mutations, possibly driven by neutral evolutionary processes, adaptive evolution, or founder effects, in BDUT 1451. Conversely, BDUT 1450 and 1452 are closely related according to cladistic analysis, consistent with the overall low evolutionary rates observed in the matK region. Overall, the analysis of the matK region revealed high conservation with subtle divergence among the three cultivars.

The estimated shape parameter of the rbcL region was 28.0140 for the discrete gamma distribution, suggesting that the heterogeneity in the rbcL region was significantly greater than that in the matK region. The mean evolutionary rates in these categories were 0.75, 0.89, 0.99, 1.09, and 1.28 substitutions per site, implying a moderate evolutionary rate with high heterogeneity. The nucleotide frequencies were A = 30.11%, T = 28.62%, C = 21.73%, and G = 19.55%, which indicated slight GC bias, contrasting with the AT bias observed in the matK analysis. The maximum log likelihood for this computation was − 2070.782. The estimated transition/transformation bias (*R*) is 0.41. The maximum log likelihood for this computation was − 2093.020. Analysis of the *rbcL* region of the tested cultivars revealed that BDUT 1451 and BDUT 1452 were closely related to each other and formed a single cluster according to both methods of phylogenetic analysis. BDUT 1450 was found in a separate cluster (Figs. 5 and 6). Both the phenetic and cladistic methods for rbcL placed BDUT 1450 in a separate cluster, while BDUT 1451 and 1452 formed a closer cluster, aligning with the matK analysis. However, the rbcL analysis revealed closer relationships between BDUT 1451 and 1452 than between BDUT 1451 and matK,

potentially indicating incomplete lineage sorting in the rbcL gene, where ancestral polymorphisms have not yet been sorted across lineages.

According to the trnH-psbA region analysis, the estimated value of the shape parameter for the discrete Gamma distribution is 2.6231. The mean evolutionary rates in these categories were 0.32, 0.61, 0.88, 1.22, and 1.97 substitutions per site. These results suggest moderate rate variation across sites in the trnH-psbA region. In addition, these findings indicate a mix of conserved and faster-evolving regions compared to the matK and rbcL barcode regions. The nucleotide frequencies were A = 31.48%, T = 33.73%, C = 17.95%, and G = 16.84%, which indicated a rich AT-bias similar to that of matK, which features characteristic chloroplast DNA. The maximum log likelihood for this computation was −8919.082. The estimated R bias is 0.68, which is comparatively greater than that of matK and rbcL, potentially reflecting differences in structural constraints in this region. The maximum log likelihood for R bias computation was − 9236.580. In the phenetic NJ and cladistic MP methods of phylogenetic analysis of *trnH-psbA*, BDUT 1451 and BDUT 1452 were linked in a single cluster, and this cluster diverged from the cultivar BDUT 1450 (Figs. 7 and 8). The overall pairwise distance among the cultivars was 1.289 for the *trnH–pspA* sequence. This strengthens the confidence in the observed divergence and suggests a fundamental split between BDUT 1450 and the other two cultivars. The closer relationship between BDUT 1451 and BDUT 1452 in rbcL and trnH-psbA, compared to the more distinct separation in matK, suggested potential incomplete lineage sorting (ILS) in these regions.

According to the results of the trnL sequence analysis, the estimated shape parameter for the discrete Gamma distribution was 28.2892, similar to that of rbcL, indicating considerable rate of variation across sites. Similarly, the mean evolutionary rates in these categories resembled those of the rbcL region and were 0.75, 0.89, 0.99, 1.09, and 1.28 substitutions per site, suggesting comparable evolutionary dynamics in these regions. The nucleotide frequencies were A = 37.97%, T = 27.16%, C = 15.57%, and G = 19.31%, similar to those of matK and trnH-psbA. The maximum log likelihood for this computation was −1030.978. The estimated transition/transformation bias (*R*) is 0.53. The maximum log likelihood for transition/transformation computation was − 1074.046. The *trnL* sequence analysis showed that BDUT 1450 and 1452 formed a single cluster, and BDUT 1451 was found in a separate cluster via the phenetic neighbor–joining (NJ) method and the cladistic difference–platelet method. Similarly, BDUT 1452 diverged from the BDUT 1451 cluster, and BDUT 1450 diverged from BDUT 1452 (Figs. 9 and 10). The overall pairwise distance between them was 1.845. Thus, trnL uniquely clusters BDUT 1450 and 1452 together, contrasting with the groupings observed for matK, rbcL, and trnH-psbA.

Species discrimination was successful in an earlier study with 72% of the cases, with the remaining species being matched to groups of congeneric species with 100% success using rbcL and matK. According to the trnH-psbA and rbcL sequence analyses, BDUT 1451 and BDUT 1452 were closely related to each other. BDUT 1450 evolved from a separate clade with *A. textile* as a closely related species according to the cladistic method and diverged from a cluster with *A. textile* JX848396 and *A. sikkimense* isolate LHXJ0489 according to the phenetic method of *rbcL* analysis. According to the phenetic method, for *trnL*, BDUT 1452 and BDUT 1450 were found in a single cluster, which showed that

they were closely related, and BDUT 1451 was found in a separate clade related to *A. linearifolium* JF262666; therefore, these genes diverged from a large clade with different *Allium* species.

The ISSR primers used in this study produced a total of 91 scorable bands with sizes ranging from ~ 250 bp to 1500 bp (Figs. 11–13). The total number of bands for each primer ranged from zero to 21, with an average of 13 (Table 3). The total number of polymorphic alleles and the percentage of polymorphisms were 11 and 21.09%, respectively. This implies that low genetic diversity and limited polymorphism occur among the three cultivars. The ISSR9 and ISHY4 primers did not produce any scorable fragments in BDUT 1450, and ISSRa was observed to be an inefficient primer for amplifying all three garlic clones. ISHY 1b and ISHY2 produced the maximum number of amplified fragments (21 bands).

Table 3
Characteristics of the ISSR primers

Primer	Primer sequences 5'-3'	Annealing Temperature (°C)	AF	PF	% P
ISSR8US	G(AACA)4	50	14	4	28.57
ISSR9	HVH(TC)8	50	11	3	27.27
ISHY 1b	(GA)8YT	39	21	3	14.28
ISHY 2	(AC)8YG	43	21	4	19.0
ISHY 3	(AG)8YT	39	19	3	15.79
ISHY 4	(CA)8RG	44	5	1	20.00
ISSR a	GCV(TG)8	51	-	-	-

V = G; A; C H = A; C D + G; A; T B = G; C; T Y = C; T R = A;G

AF-Total amplified fragments; PF- Number of polymorphic amplicons; % P- Percentage of polymorphism

The highest percentage of polymorphisms was observed for ISSR8US, followed by ISSR9 (28.57% and 27.27%), suggesting that these regions offer potentially informative markers for differentiating these cultivars. The statistics of genetic variation for all polymorphic loci were 2 for the observed number of alleles (na), 1.6831 ± 0.22 for the effective number of alleles (ne), 0.3969 ± 0.79 for Nei's gene diversity (h) (1972) and 0.5837 ± 0.086 for Shannon's information index (I). These results indicate moderate genetic variation across the analysed loci. In this study, a UPGMA dendrogram of garlic cultivars was constructed on the basis of ISSR data obtained by using POPGENE *v32* software; BDUT 1450 and BDUT 1452 were linked together in a single cluster, whereas BDUT 1451 evolved as an outgroup from the BDUT 1450 and BDUT 1452 clusters. The average distance between individuals in the same cluster was determined to be 0.5862 for BDUT 1450 and BDUT 1452 and 0.5838 for BDUT 1451. This finding aligns somewhat with the clustering observed in the trnL barcoding region but contrasts with the consistent

grouping in the matK, rbcL, and trnH-psbA regions. The estimated Ln probability of the data was − 10.3, the mean value of the maximum likelihood was − 9.8, the variance in the maximum likelihood was 1.0, and the mean value of the alpha was 0.3288. The Bayesian proportion of individual plants for a K = 3 population model was identified by STRUCTURE software and is indicated in Fig. 14. The proportions suggest a combination of divergence and heterogeneity among the cultivars.

Jabbes et al. (2011) revealed that, on the basis of the ISSR primers used, 7 to 21 polymorphic fragments were pragmatic, with an average of 12 fragments, which correlates with our results with the same primers used in this study. The presence of heterologous genomes was reflected by the accessibility of a comparatively high number of polymorphic ISSR markers (Jabbes et al. 2011). The results of Sadeghi and Cheghanirza (2012) demonstrated that ISSR analysis can be used not only to study the high diversity among cultivars but also to study polymorphisms. Additionally, they confirmed that ISSR techniques are advantageous for studying genetic diversity and are robust methods for characterizing plant genomes. ISSR analysis has been extensively used to determine genetic diversity in important crops and fruits and for biodiversity conservation. Furthermore, ISSRs can be used to evaluate relationships at the interspecific level (Huang and Sun 2000). Hao et al. (2002) provided a clear picture of the relationships among closely related congeneric species. Jabbes et al. (2011) reported that the diversity between accessions is always greater than the diversity within accessions.

The discriminative ability of each marker was evaluated by the diversity index value. A higher diversity index indicates more marker information (Kandasamy et al. 2013). The authors concluded that the ISSR marker system was highly informative and efficient for observing the genetic diversity of flowering plants via analytical methods. A total of 99.47% of the bands obtained in the Mukherjee et al. (2013) study were polymorphic. Nagaoka and Ogihara (1997) reported that ISSR primers produce reliable and reproducible bands.

Conclusion

The DNA barcoding results of this study provided us with insight into the genetic diversity of three selected cultivars of *Allium sativum*. Among the cultivars, BDUT 1450 and 1452 seemed to be closely related, while BDUT 1451 diverged significantly. Discrepancies in clustering patterns among different markers suggest incomplete lineage sorting (ILS), where ancestral traits have not been fully separated. Furthermore, moderate interspecific diversity and intraspecific divergence were observed with complex genetic relationships. However, further studies with advanced methods and additional information from various cultivars from different geographical locations are needed to fully understand the evolutionary history of this important aromatic cum nutraceutical crop. Understanding this genetic landscape presents exciting opportunities for garlic breeding programs, allowing for the development of cultivars with enhanced nutritional profiles, improved disease resistance, and improved adaptation to specific environments.

Declarations

Funding & Competing Interests

The authors have no competing interests to declare that they are relevant to the content of this article.

Author Contributions

All the authors substantially contributed to the conception and design of the project. Packia Lekshmi and Rajesh produced, assembled and analysed the data. The manuscript was interpreted and written by Packia Lekshmi and Mahamuni, with valuable contributions from Rajesh and Brindha. All the authors reviewed and approved the current version of the manuscript.

Data Availability

GenBank Accession Numbers: KF769491, KF769492, KF769493, KF769497, KF769498, KF769499, KF769503, KF769504, KF769505, KF769509, KF769510, KF769511, KF779159, KF779160, KF779161.

Acknowledgements

The first author gratefully acknowledges Dr. M.B. Viswanathan, former Professor and Head, Department of Botany, Bharathidasan University, for providing the laboratory facility for DNA and PCR analyses, and all the authors gratefully acknowledge Mr. R. Anand for his assistance in sample collection.

References

1. Appeltans W, Ahyong ST, Anderson G, Angel MV, Artois T, Bailly N, Bamber R, Barber A, Bartsch I, Berta A, Błażewicz-Paszkowycz M (2012) The magnitude of global marine species diversity. Curr Biol 22:2189-2202. https://doi.org/10.1016/j.cub.2012.09.036
2. Aryakia E, Karimi HR, Naghavi MR et al. (2016) Morphological characterization of intra-and interspecific diversity in some Iranian wild *Allium* species. Euphytica 211:185-200. https://doi.org/10.1007/s10681-016-1729-8
3. Asahina H, Shinozaki J, Masuda K, Morimitsu Y, Satake M (2010) Identification of medicinal Dendrobium species by phylogenetic analyses using *matK* and *rbcL* sequences. J Nat Med 64:133-138. https://doi.org/10.1007/s11418-009-0379-8
4. Bafeel SO, Arif IA, Bakir MA, Khan HA, Farhan AHA, Homaidan AAA, Ahamed A, Thomas J (2011) Comparative evaluation of PCR success with universal primers of maturase K (matK) and ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (rbcL) for barcoding of some arid plants. Plant Omics J 4:195-198 https://doi.org/10.3316/informit.310811888981920
5. Baghalian K, Naghavi MR, Ziai SA, Badi HN (2006) Post-planting evaluation of morphological characters and allicin content in Iranian garlic (*Allium sativum* L.) ecotypes. Sci Hort 107:405-410. https://doi.org/10.1016/j.scienta.2005.11.008
6. Baldwin BG, Sanderson MJ, Porter JM, Wojciechowski MF, Campbell CS, Donoghue MJ (1995) The ITS region of nuclear ribosomal DNA: A valuable source of evidence on angiosperm phylogeny. Ann

Missouri Bot Gard 82:247-277. https://doi.org/10.2307/2399880

7. Besse P, Da Silva D, Grisoni M (2021) Plant DNA Barcoding Principles and Limits: A Case Study in the Genus *Vanilla*. *In: Molecular Plant Taxonomy. Methods in Molecular Biology*, vol 2222. Humana, New York, NY. pp.131-148. https://doi.org/10.1007/978-1-0716-0997-2_8
8. Bieniek W, Mizianty M, Szklarczyk M (2015) Sequence variation at the three chloroplast loci (matK, rbcL, trnH-psbA) in the Triticeae tribe (Poaceae): comments on the relationships and utility in DNA barcoding of selected species. Plant Syst Evol 301:1275-1286. https://doi.org/10.1007/s00606-014-1138-1
9. Blattner FR, Kadereit JW (1999) Morphological evolution and ecological diversification of the forest dwelling poppies (Papaveracea: Chelidonioideae) as deduced from a molecular phylogeny of the ITS region. Pl Syst Evol 219:181-197. https://doi.org/10.1007/BF00985578
10. Burgess KS, Fazekas AJ, Kesanakurti PR, Graham SW, Husband BC,Newmaster SG, Percy DM, Hajibabaei M, Barrett SCH (2011) Discriminating plant species in a local temperate flora using the *rbcL+matK*DNA barcode. Method Ecol Evol 2:333-340. https://doi.org/10.1111/j.2041-210X.2011.00092.x
11. CBOL Plant Working Group (2009) A DNA barcode for land plants. Proc Natl Acad Sci USA 106:12794-12797. https://doi.org/10.1073/pnas.090584510
12. Cetiz MV, Turumtay EA, Burnaz NA, Özhatay FN, Kaya E, Memon A, Turumtay H (2023) Phylogenetic analysis based on the ITS, mat K and rbc L DNA barcodes and comparison of chemical contents of twelve Paeonia taxa in Türkiye. Mol Biol Rep 30:1-4. https://doi.org/10.1007/s11033-023-08435-z
13. Chac LD, Thinh BB (2023) Species identification through DNA barcoding and its applications: A review. Biol Bull Russ Acad Sci 50:1143-1156. https://doi.org/10.1134/S106235902360229X
14. Chen SL, Yao H, Han JP et al. (2010) Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. PLoS One 5:e8613. https://doi.org/10.1371/journal.pone.0008613
15. Costa FO, Carvalho GR (2010) New insights into molecular evolution: prospects from the Barcode of Life Initiative (BOLI). Theory Biosci 129:149-157. https://doi.org/10.1007/s12064-010-0091-y
16. Costello MJ, Wilson S, Houlding B (2013) More taxonomists describing significantly fewer species per unit effort may indicate that most species have been discovered. Syst Biol 62:616-624. https://doi.org/10.1093/sysbio/syt024
17. Costion C, Ford A, Cross H, Crayn D, Harrington M, Lowe A (2011) Plant DNA Barcodes Can Accurately Estimate Species Richness in Poorly Known Floras. PLoS One 6:e26841. https://doi.org/10.1371/journal.pone.0026841
18. Creer S, Deiner K, Frey S, Porazinska D, Taberlet P, Thomas WK, Potter C, Bik HM (2016) The ecologist's field guide to sequence-based identification of biodiversity. Methods Ecol Evol 7:1008-1018. https://doi.org/10.1111/2041-210X.12574
19. Deiner K, Bik HM, Mächler E, Seymour M, Lacoursière-Roussel A, Altermatt F, Creer S, Bista I, Lodge DM, De Vere N, Pfrender ME (2017) Environmental DNA metabarcoding: Transforming how we

survey animal and plant communities. Mol. Ecol. 26:5872-5895. https://doi.org/10.1111/mec.14350

20. DeSalle R, Goldstein P (2019) Review and interpretation of trends in DNA barcoding. Front Ecol Evol 7:302. https://doi.org/10.3389/fevo.2019.00302
21. Dincă V, Dapporto L, Somervuo P, Vodă R, Cuvelier S, Gascoigne-Pees M, Huemer P, Mutanen M, Hebert PD, Vila R (2021) High resolution DNA barcode library for European butterflies reveals continental patterns of mitochondrial genetic diversity. Commun. Biol 4:315. https://doi.org/10.1038/s42003-021-01834-7
22. Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull 19:11-15
23. Dubouzet JG, Shinoda K (1999) Relationships among old and new world *Alliums* according to ITS DNA sequence analysis. Theor Appl Genet 98:422-433. https://doi.org/10.1007/s001220051088
24. Ebach MC, Holdrege C (2005) DNA barcoding is no substitute for taxonomy. Nature 434:697. https://doi.org/10.1038/434697b
25. Engel MS, Ceríaco LM, Daniel GM et al (2021) The taxonomic impediment: a shortage of taxonomists, not the lack of technical approaches. Zool J Linn Soc 193:381-387. https://doi.org/10.1093/zoolinnean/zlab072
26. Fernandes TJ, Amaral JS, Mafra I (2021) DNA barcode markers applied to seafood authentication: An updated review. Crit Rev Food Sci Nutr 61:3904-3935. https://doi.org/10.1080/10408398.2020.1811200
27. 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. https://doi.org/10.1023/A:1012069532157
28. Figliuolo G, Di Stefano D (2007) Is single bulb producing garlic *Allium sativum* or *Allium ampeloprasum*?. Sci Hort 114:243-249. https://doi.org/10.1016/j.scienta.2007.06.021
29. Fredotović Z, Šamanić I, Weiss-Schneeweiss H, Kamenjarin J, Jang TS, Puizina J (2014) Triparental origin of triploid onion, *Allium* × *cornutum* (Clementi ex Visiani, 1842), as cytogenetic analyses. BMC Plant Biol 14:24. https://doi.org/10.1186/1471-2229-14-24
30. Friesen N, Fritsch RM, Pollner S, Blattner FR (2000) Molecular and morphological evidence for an origin of the aberrant genus *Milula* within Himalayan species of *Allium* (Amaryllidaceae). Mol Phylogenet Evol 17:209-218. https://doi.org/10.1006/mpev.2000.0844
31. Fu YM, Jiang WM, Fu CX (2011) Identification of species within *Tetrastigma* (Miq.) Planch. (Vitaceae) based on DNA barcoding techniques. J Syst Evol 49(3):237-245. https://doi.org/10.1111/j.1759-6831.2011.00126.x
32. Gao T, Yao H, Song J, Zhu Y, Liu C, Chen S (2010) Evaluating the feasibility of using candidate DNA barcodes in discriminating species of the large Asteraceae family. BMC Evol Biol 10:324-330. https://doi.org/10.1186/1471-2148-10-324
33. Garg S, Biju SD (2021) DNA barcoding and systematic review of Minervaryan frogs (Dicroglossidae: Minervaryd) of peninsular India: Resolution of a taxonomic conundrum with description of a new

species. Asian Herpetol Res 12:345-378. https://doi.org/ 10.16373/j.cnki.ahr.210023

34. Gostel MR, Kress WJ (2022) The expanding role of DNA barcodes: Indispensable tools for ecology, evolution, and conservation. Diversity 14:213. https://doi.org/10.3390/d14030213
35. Hao G, Lee DH, Lee JS, Lee NS (2002) A study of taxonomical relationship among species of Korean *Allium* sect. *Sacculiferum* (Amaryllidaceae) and related species using inter simple sequence repeats (ISSR) markers. Bot Bull Acad Sin 43:63-68
36. Hebert PDN, Alina Cywinska, Ball SJ, deWaard JR (2003) Biological identifications through DNA barcodes. Proc R Soc Lond B 270:313-321. https://doi.org/10.1098/rspb.2002.2218
37. Hebert PDN, Gregory DTR (2005) The Promise of DNA Barcoding for Taxonomy. Syst Biol 54:852-859. https://doi.org/10.1080/10635150500354886
38. Hirschegger P, Jaske J, Trontely P, Bohanec B (2010) Origins of *Allium ampeloprasum* horticultural groups and a molecular phylogeny of the section *Allium* (*Allium*: Amaryllidaceae). Mol Phylogenet Evol 54:488-497. https://doi.org/10.1016/j.ympev.2009.08.030
39. Ho VT, Tran TK, Vu TT, Widiarsih S (2021) Comparison of matK and rbcL DNA barcodes for genetic classification of jewel orchid accessions in Vietnam. J Genet Eng Biotechnol 19:93. https://doi.org/10.1186/s43141-021-00188-1
40. Honeycutt RL (2021) DNA barcodes: Controversies, mechanisms, and future applications. Front Ecol Evol 9:718865. https://doi.org/10.3389/fevo.2021.718865
41. Hsiao C, Jacobs SWL, Chatterton NJ, Asay KH (1999) A molecular phylogeny of the grass family (Poaceae) based on the sequence of nuclear ribosomal DNA (ITS). Aust Syst Bot 11:667-688
42. Huang JC, Sun M (2000) Genetic diversity and relationships of sweet potato and its wild relatives in *Ipomeaser* is Batatas (Convolvulaceae) as revealed by Inter Simple Sequence Repeat (ISSR) and restriction analysis of chloroplast DNA. Theor Appl Genet 100:1050-1060. https://doi.org/10.1007/s001220051386
43. iBOL, The International Barcode of Life Consortium (2024) International Barcode of Life project (iBOL). https://doi.org/10.15468/inygc6. Accessed 07 January 2024
44. Intharuksa A, Sasaki Y, Ando H, Charoensup W, Suksathan R, Kertsawang K, Sirisa-Ard P, Mikage M (2020) The combination of ITS2 and psb A-trn H region is powerful DNA barcode markers for authentication of medicinal Terminalia plants from Thailand. J Nat Med 74:282-293. https://doi.org/10.1007/s11418-019-01365-w
45. Ipek M, Ipek A, Simon PW (2008) Rapid characterization of garlic clones with Locus-specific DNA markers. Turkis J Agric Forestry 32:357-3622
46. Jabbes N, Geoffriau EE, Le Clerc V, Dridi B, Hannechi C (2011) Inter simple sequence repeat fingerprints for assess genetic diversity of Tunisian garlic populations. J Agric Sci 3:77-85. https://doi.org/10.5539/jas.v3n4p77
47. Jarman SN, Elliott NG (2000) DNA evidence for morphological and cryptic Cenozoicspeciations in the Anaspididae,'living fossils' from the Triassic. J Evol Biol 13:624–633. https://doi/abs/10.1046/j.1420-9101.2000.00207.x

48. Jiang Z, Zhang M, Kong L, Bao Y, Ren W, Li H, Liu X, Wang Z, Ma W (2023) Identification of Apiaceae using ITS, ITS2 and psba-trnH barcodes. Mol Biol Rep 50:245-253. https://doi.org/10.1007/s11033-022-07909-w
49. Kandasamy T, Kumar K, Kaprakkaden A, Lohet VD, Gosh J (2013) Molecular diversity analysis of flower colour variants of *Butea monosperma* (Lam.) TAUB. using Inter Simple Sequence Repeats. The Biscan 8:967-974
50. Kimura M (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J Mol Evol 16:111-120. https://doi.org/10.1007/BF01731581
51. Knowlton N (1993) Sibling species in the sea. Ann Rev Ecol Syst 24:189-216. https://doi.org/10.1146/annurev.es.24.110193.001201
52. Kress WJ, Wurdack KJ, Zimmer EA, Weigt LA, Janzen DH (2005) Use of DNA barcodes to identify flowering plants. Proc Natl Acad Sci *USA* 102:8369-8374. https://doi.org/10.1073/pnas.0503123102
53. Kuo LY, Li FW, Chiou WL, Wang CN (2011) First insights into fern *matK* phylogeny. Mol Phylogenet Evol 59:556-566. https://doi.org/10.1016/j.ympev.2011.03.010
54. Larsen BB, Miller EC, Rhodes MK, Wiens JJ (2017) Inordinate fondness multiplied and redistributed: the number of species on earth and the new pie of life. Q Rev Biol 92:229-265. https://doi.org/10.1086/693564
55. Lewontin RC (1972) Testing the theory of natural selection. Nature 236:181-182
56. Li X, Wiens JJ (2023) Estimating global biodiversity: the role of cryptic insect species. Syst Biol 72:391-403. https://doi.org/10.1093/sysbio/syac069
57. Manzoor MT, Ali S, Fatima A, Tariq MR, Ali Q, Shafiq M, Iqbal MA, Zahid A, Hussain A (2022) Analyzing local spinach plant diversity using the rpoC1 and matK DNA barcode technique. Pak J Agri Sci 59:1029-1034. https://doi.org/10.21162/PAKJAS/22.48
58. Maulidya NN, Rohimah S, Ramadany Z, Ratnasari T, Su'udi M (2020) Assessment of the DNA Barcodes Characteristic of *Phalaenopsis deliciosa* based on matK, rbcL, and ITS. Biog J Ilm Biol 8:138-144. https://doi.org/10.24252/bio.v8i2.13278
59. McCallum J (2007) Onions. In: Kole C (ed) Genome mapping and molecular breeding in plants, Vegetable, Germany: Springer, Berlin, pp 331-347
60. Mukherjee A, Sikdar B, Ghosh B, Banerjee A, Ghosh E, Bhattachary N, Roy SC (2013) RAPD and ISSR analysis of some economically important species, varieties and cultivars of the genus *Allium* (Amaryllidaceae). Turk J Bot 37:605-618. https://doi.org/10.3906/bot-1208-18
61. Nagaoka T, Ogihara Y (1997) Applicability of inter simple sequence repeats polymorphisms in wheat for use as DNA markers in comparison to RFLP and RAPD markers. Theor Appl Genet 94:597-602. https://doi.org/10.1007/s001220050456
62. Nei M (1972) Genetic distance between populations. Amer Naturalist 106:283-292. https://doi.org/10.1086/282771

63. Nguyen NH, Driscoll HE, Specht CD (2008) A molecular phylogeny of the wild onions (*Alliums;* Amaryllidaceae) with a focus on the western North American center of diversity. Mol phylogenet Evol 47:1157-1172. https://doi.org/10.1016/j.ympev.2007.12.006
64. Noyes RD, Rieseberg LH (1999) ITS sequence data support a single origin for North American Astereae (Asteraceae) and reflect deep geographic divisions *Asters* s.l. Am J Bot 86:398-412. https://doi.org/10.2307/2656761
65. Pfeiler E (2023) Performance of DNA barcodes for informing the subspecies controversy in North American populations of Callophrys gryneus (Hübner, [1819]) (Lepidoptera: Lycaenidae). 31 Biol J Linn Soc. https://doi.org/10.1093/biolinnean/blad094
66. Piper AM, Batovska J, Cogan NO, Weiss J, Cunningham JP, Rodoni BC, Blacket MJ (2019) Prospects and challenges of implementing DNA metabarcoding for high-throughput insect surveillance. GigaScience 8:giz092. https://doi.org/10.1093/gigascience/giz092
67. Raskoti BB, Ale R (2021) DNA barcoding of medicinal orchids in Asia. Sci Rep 11:23651. https://doi.org/10.1038/s41598-021-03025-0
68. Rauf A, Abu-Izneid T, Thiruvengadam M, Imran M, Olatunde A, Shariati MA, Bawazeer S, Naz S, Shirooie S, Sanches-Silva A, Farooq U (2022) Garlic (*Allium sativum* L.): Its chemistry, nutritional composition, toxicity, and anticancer properties. Curr Top Med Chem 22:957-972. https://doi.org/10.2174/1568026621666211105094939
69. Sadeghi A, Cheghairza K (2012) Efficiency of RAPD and ISSR markers system for studying genetic diversity in common bean *Phaseolus rulgarish* L. cultivars. Ann Biol Res 3:3267-3273
70. Sawarkar AD, Shrimankar DD, Kumar M, Kumar P, Kumar S, Singh L (2021) Traditional system versus DNA barcoding in identification of bamboo species: a systematic review. Mol Biotechnol 63:651-675. https://doi.org/10.1007/s12033-021-00337-4
71. Starr JR, Naczi RF, Chouinard BN (2009) Plant DNA barcodes and species resolution in sedges (Carex, Cyperaceae). Mol Ecol Resour Suppl s1:151-163. https://doi.org/10.1111/j.1755-0998.2009.02640.x
72. Taberlet P, Gielly L, Pautou G, Bouvet J (1991) Universal primers for amplification of three non-coding regions of chloroplast DNA. Pl Mol Biol 17:1105-1109. https://doi.org/10.1007/BF00037152
73. Tamura K, Nei M (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol Biol Evol 10:512-526
74. Tekle YI, Wood FC (2018) A practical implementation of large transcriptomic data analysis to resolve cryptic species diversity problems in microbial eukaryotes. BMC Evol Biol 18:170. https://doi.org/10.1186/s12862-018-1283-1
75. Tollefson J (2019). One million species face extinction. Nature 569:171. https://doi.org/10.1038/d41586-019-01448-4
76. 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. https://doi.org/10.1007/s10681-014-1097-1

77. White TJ, Bruns T, Lee S, Taylor J (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetids. In: Innis MA, Gelfand Dh, Sninsky JJ, White TJ (eds) PCR protocols: a guide to methods and applications. Academic Press, New York, pp 315-322
78. Xie DF, Tan JB, Yu Y, Gui LJ, Su DM, Zhou SD, He XJ (2020) Insights into phylogeny, age and evolution of *Allium* (Amaryllidaceae) based on the whole plastome sequences. Ann Bot 125:1039-1055. https://doi.org/10.1093/aob/mcaa024
79. Yang CQ, Lv Q, Zhang AB (2020) Sixteen years of DNA barcoding in China: What has been done? What can be done? Front Ecol Evol 8:57. https://doi.org/10.3389/fevo.2020.00057
80. Zhang X, Guo F, Huang X et al (2024) Combined physiological and transcriptomic analyses to identify candidate genes involved in aging during storage of *Allium mongolicum* Regel. seeds. Euphytica 220:14. https://doi.org/10.1007/s10681-023-03259-1

Figures

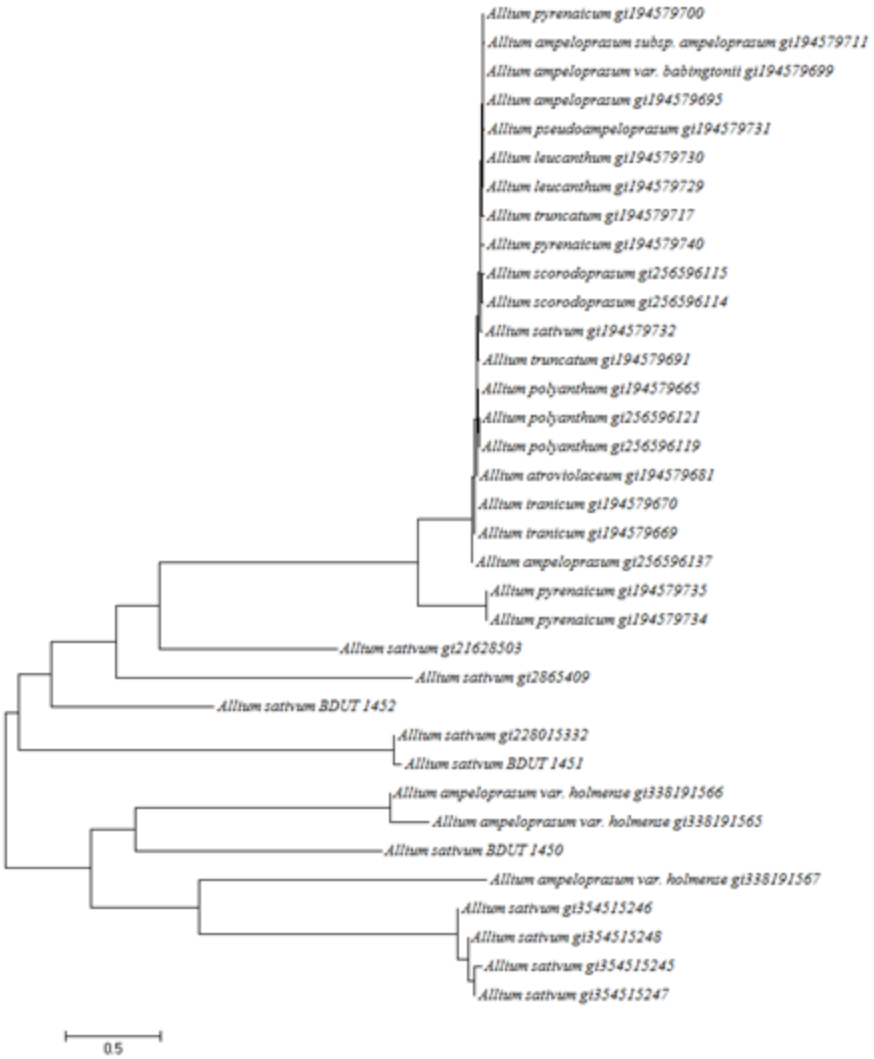

Figure 1

NJ (phenetic method) tree based on the ITS region of *Allium sativum*

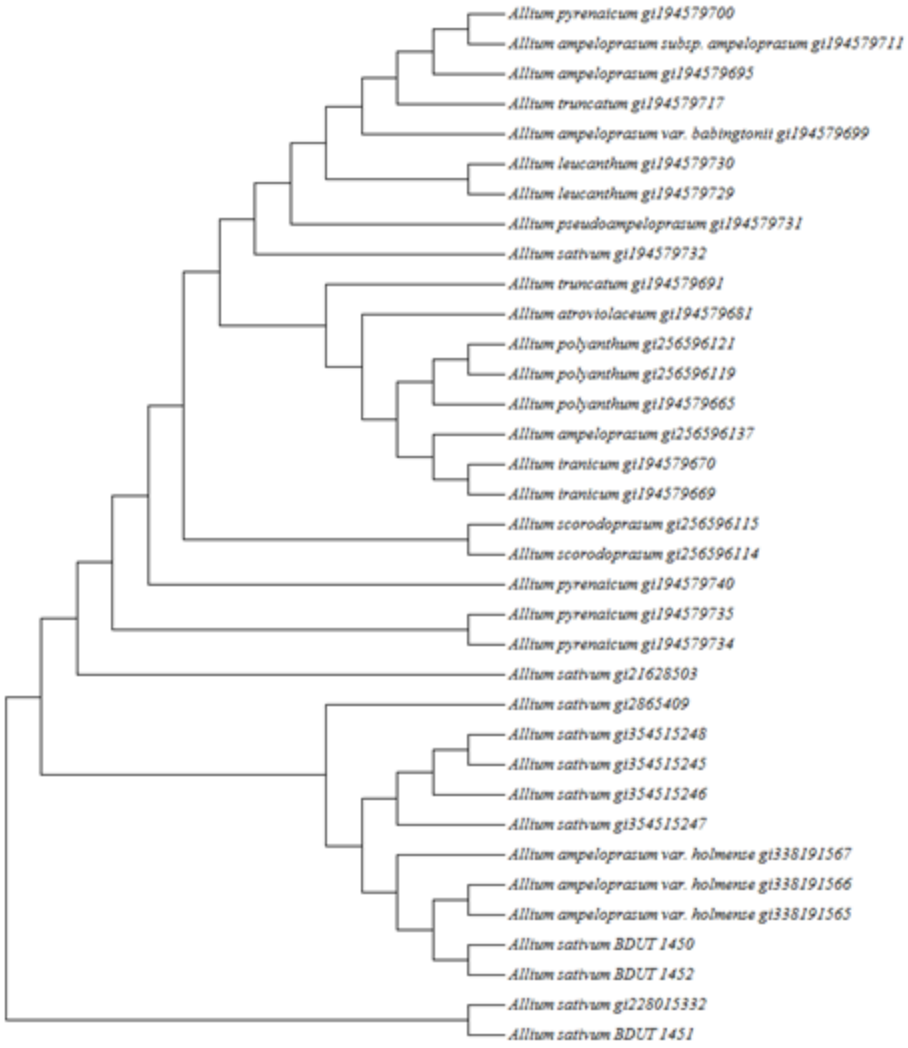

Figure 2

Cladistic method (MP) tree based on the ITS region of *Allium sativum*

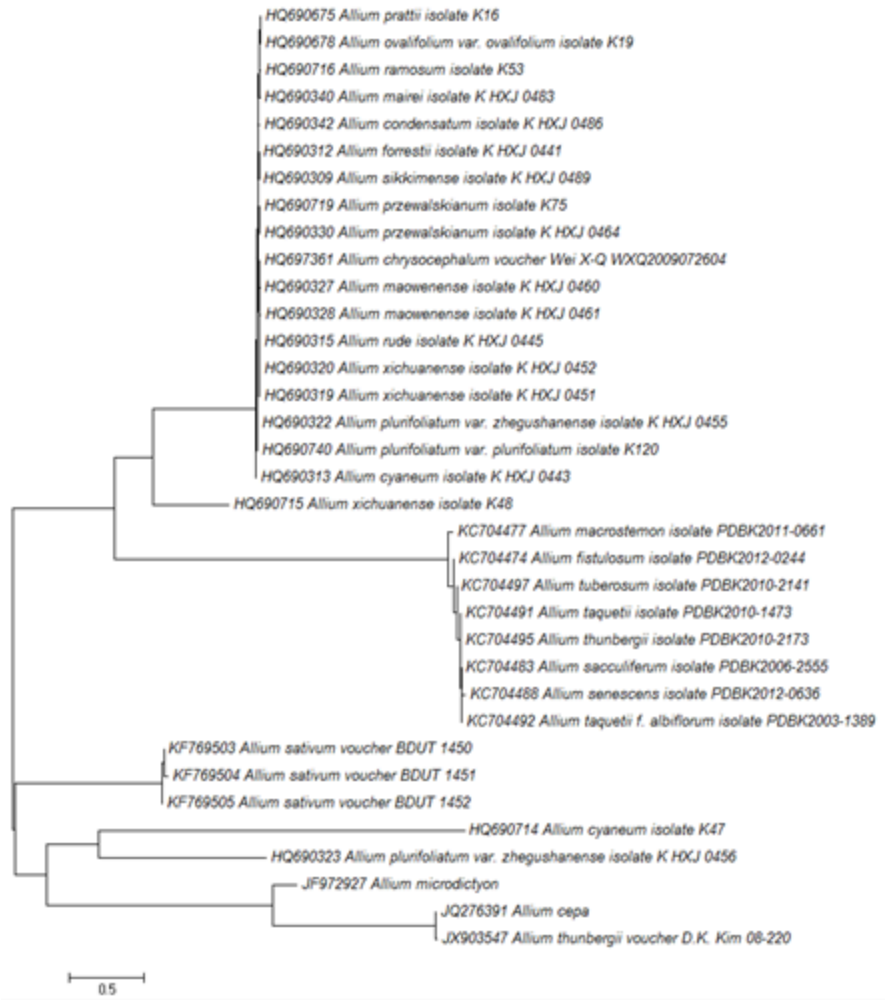

Figure 3

NJ (phenetic method) tree based on the *matK* region of *Allium sativum*

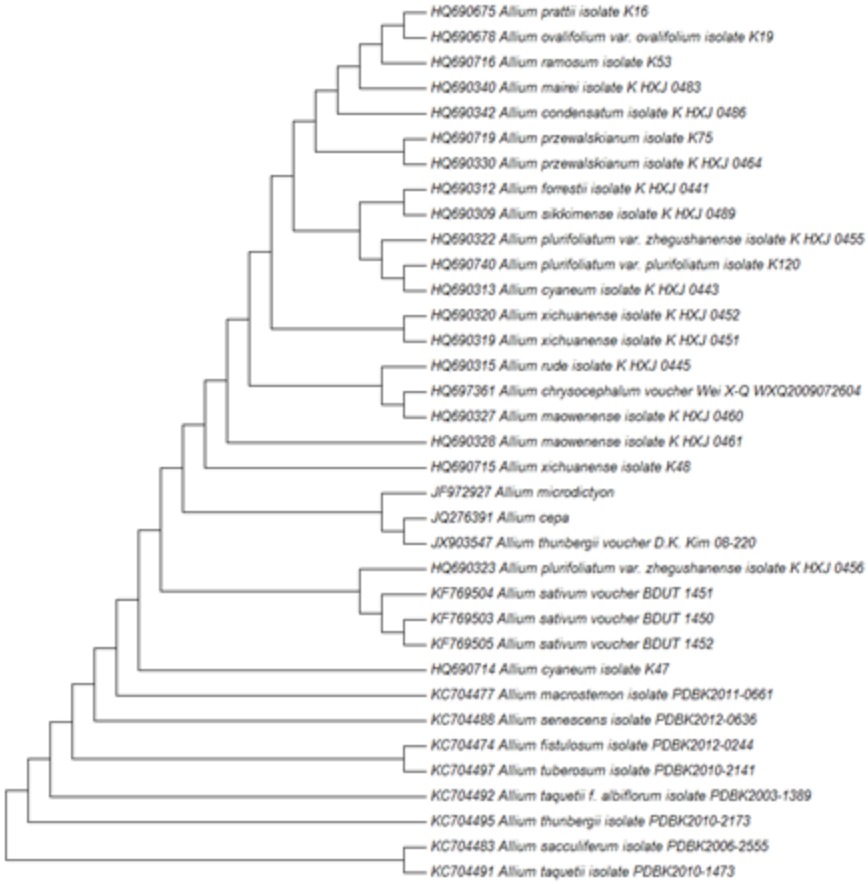

Figure 4

Cladistic method (MP) tree based on the *matK* region of *Allium sativum*

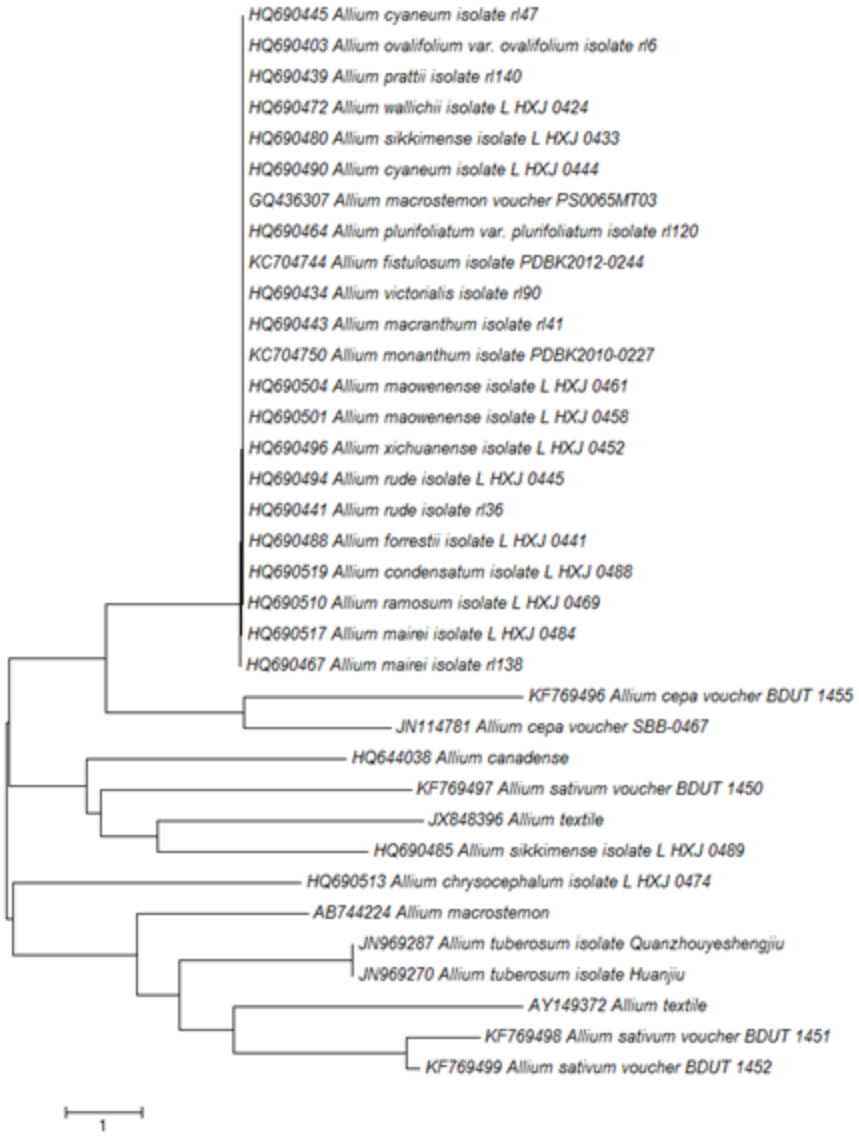

Figure 5

NJ (phenetic method) tree based on the *rbcL* region of *Allium sativum*

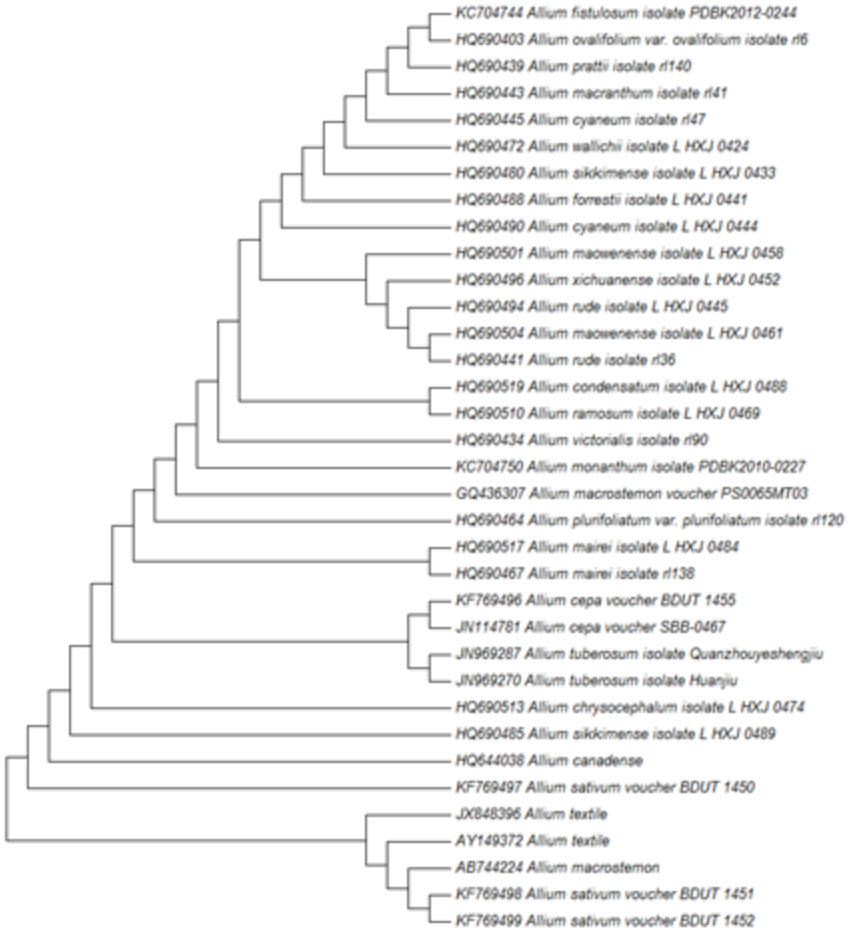

Figure 6

Cladistic method (MP) tree based on the *rbcL* region of *Allium sativum*

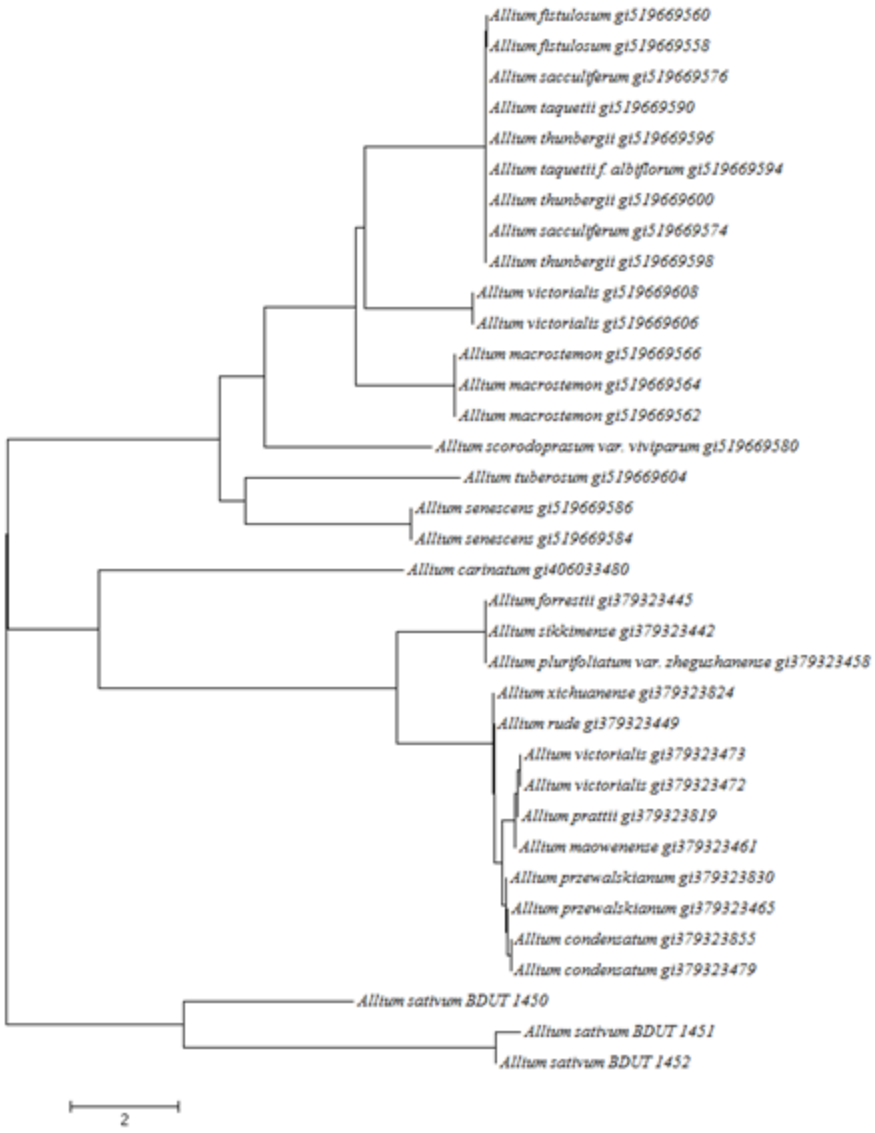

Figure 7

Phylogenetic method (NJ) tree based on the *trnH-psbA* region of *Allium sativum*

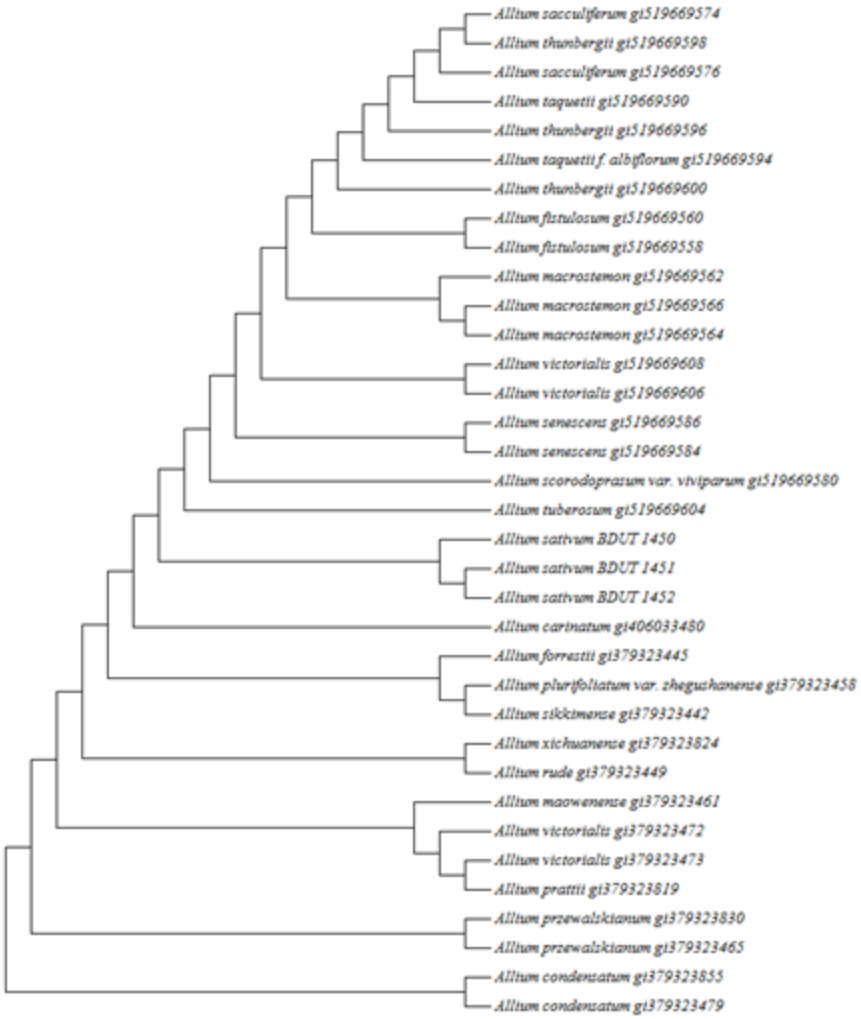

Figure 8

Cladistic method (MP) tree based on the *trnH-psbA* region of *Allium sativum*

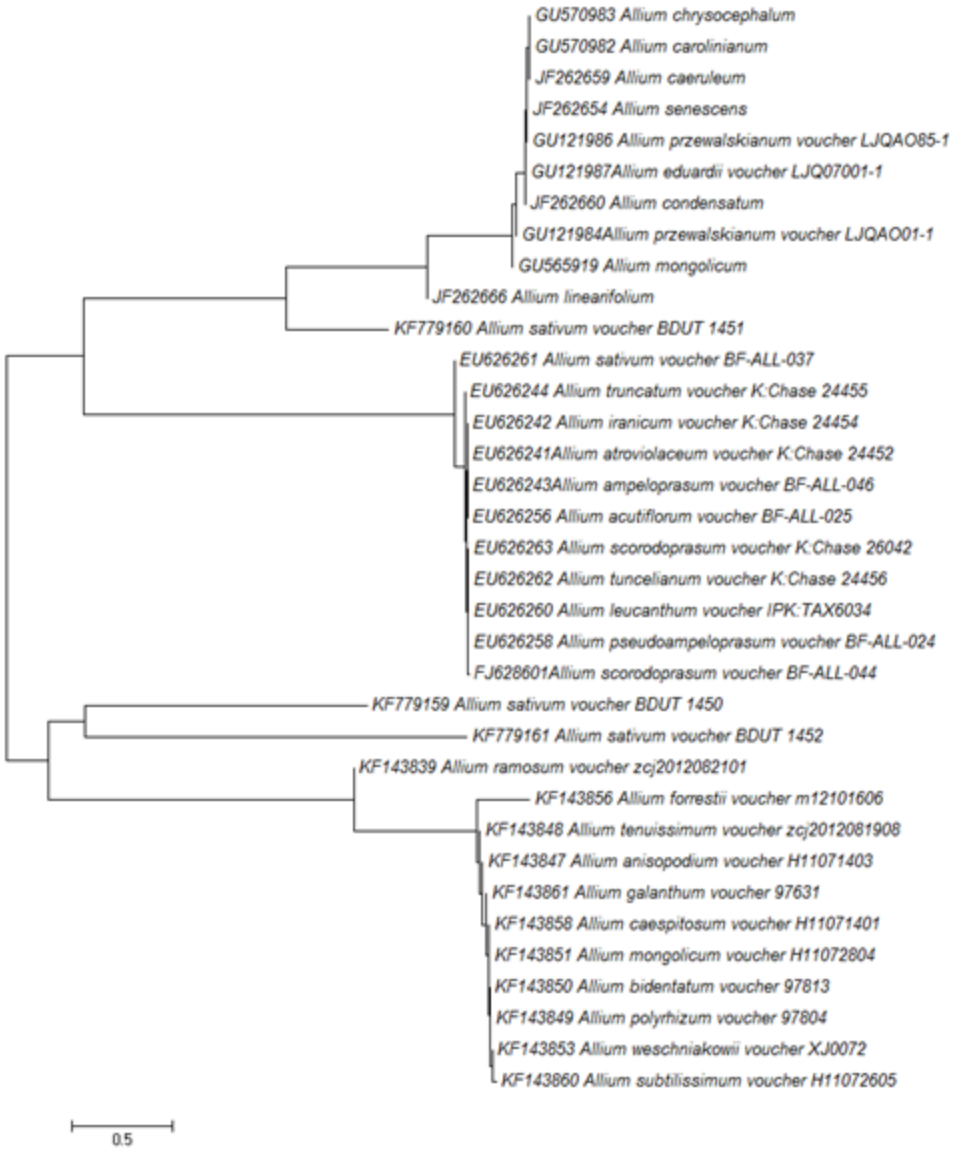

Figure 9

Phylogenetic method (NJ) tree based on the *trnL* region of *Allium sativum*

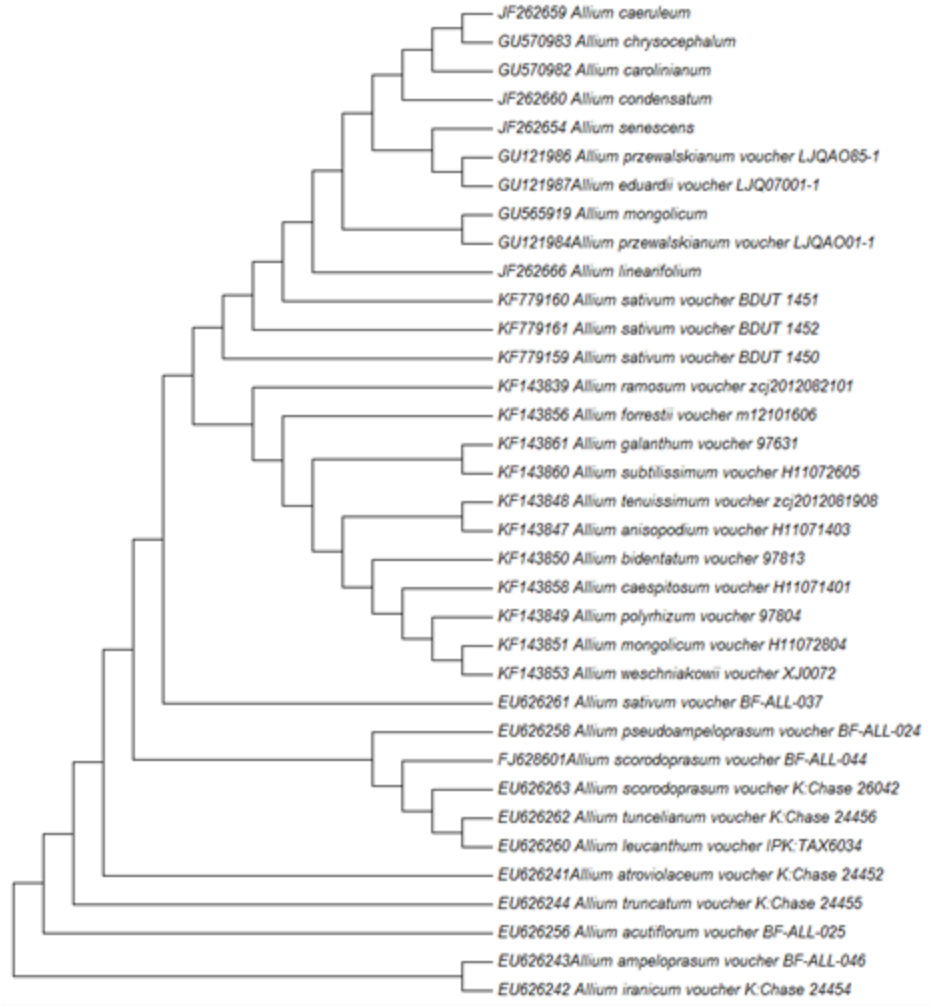

Figure 10

Cladistic method (MP) tree based on the *trnL* region of *Allium sativum*

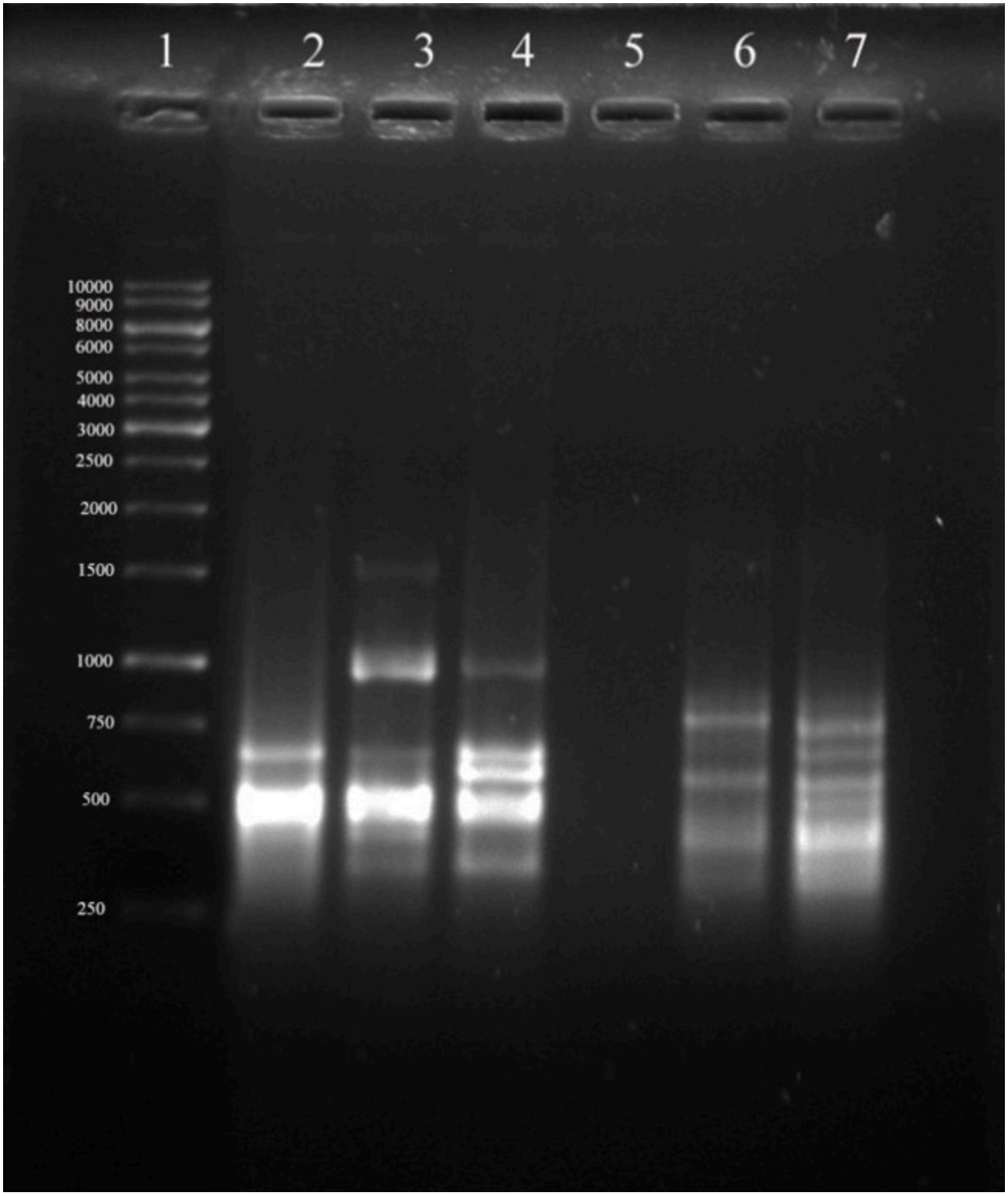

Figure 11

Amplified products of garlic cultivars for the ISSR8US and ISSR9 primers

(Lane 1- DNA ladder; Lanes 2, 3 and 4- Amplified products of BDUT 1450, BDUT 1451 and BDUT 1452 for the primer pair ISSR8US; Lanes 5, 6 and 7- Amplified products of BDUT 1450, BDUT 1451 and BDUT 1452 for the primer pair ISSR9)

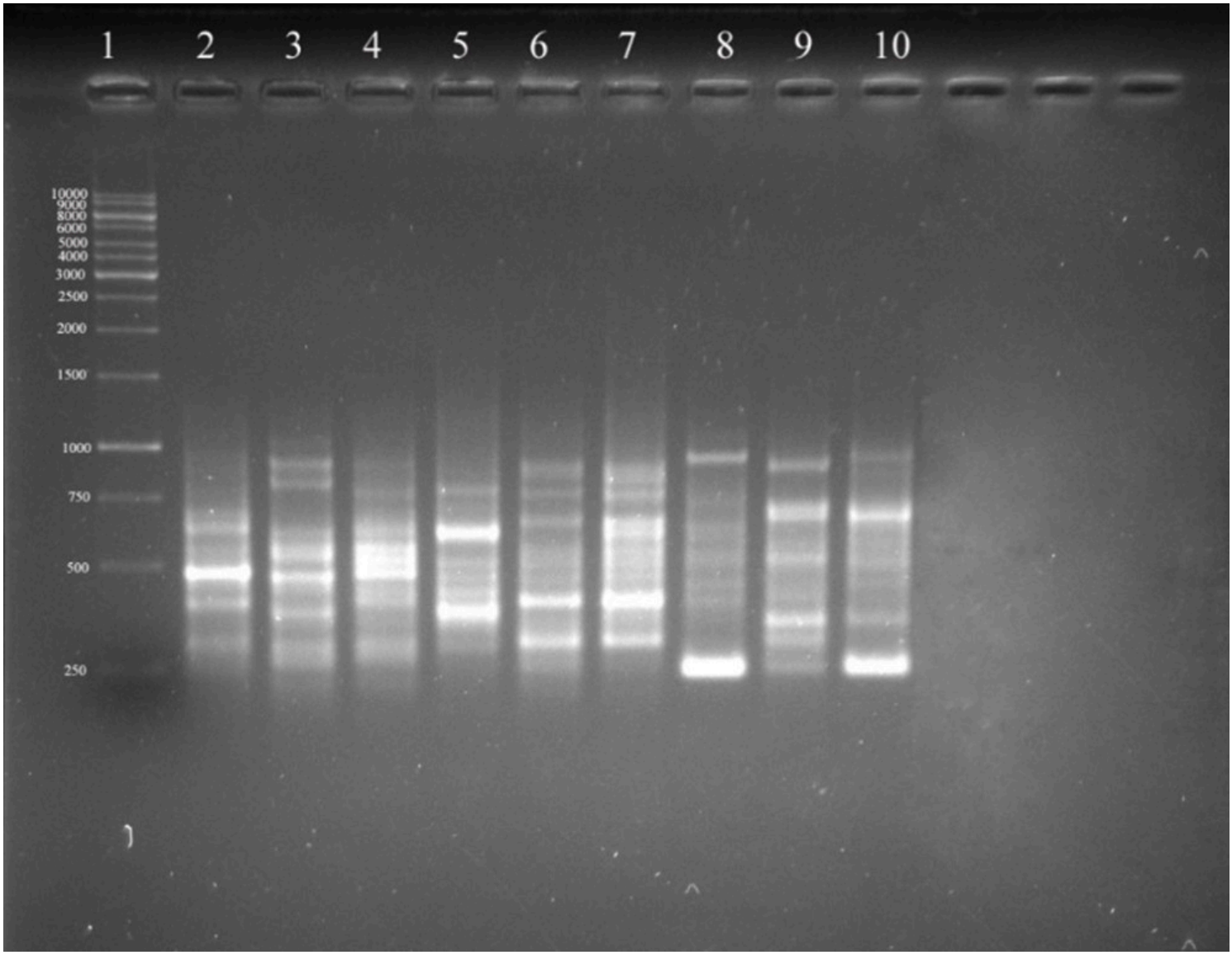

Figure 12

Amplified products of garlic cultivars for ISHY 1b, ISHY 2 and ISHY 3 primers

(Lane 1- DNA ladder; Lanes 2, 3 and 4 – Amplified products of BDUT 1450, BDUT 1451 and BDUT 1452 for the primer ISHY 1b; Lanes 5, 6 and 7 – Amplified products of BDUT 1450, BDUT 1451 and BDUT 1452 for the primer ISHY 2; Lanes 8, 9 and 10 – Amplified products of BDUT 1450, BDUT 1451 and BDUT 1452 for the primer ISHY 3)

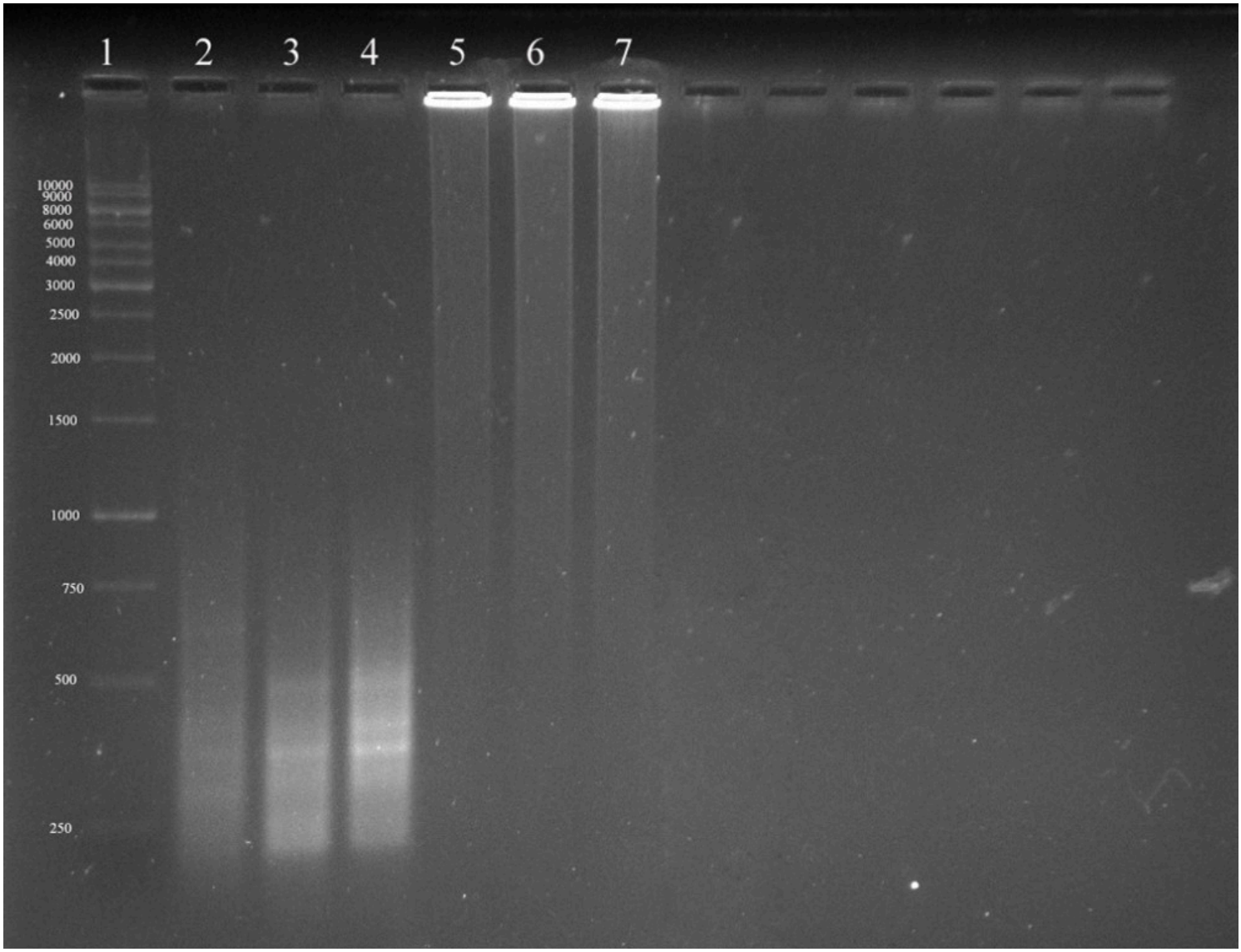

Figure 13

Amplified products of garlic cultivars for ISHY4 and ISSRa primers

Lane 1- DNA ladder

Lanes 2, 3 and 4 – Amplified products of BDUT 1450, BDUT 1451 and BDUT 1452, respectively, for the primer ISHY 4

Lanes 5, 6 and 7 – Amplified products of BDUT 1450, BDUT 1451 and BDUT 1452 for the primer pair ISSR a

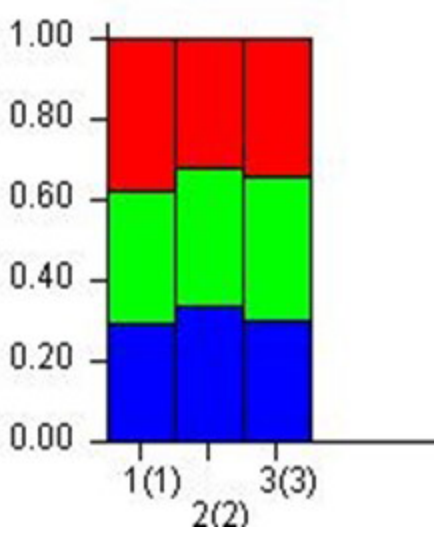

Figure 14

Bayesian proportions of individual plants for the K=3 population model. The populations identified by STRUCTURE software are indicated in different colours.