*OnLine Journal of Biological Sciences*

Original Research Paper

# Wild *Allium longicuspis* Regel is a Feral Form of *Allium sativum* L. in Kazakhstan: A Comparative Molecular Genetic Analysis

**[1,2]Akerke Tokenova, [2]Gulnar Sitpayeva, [2]Nadezhda Gemejiyeva, [1]Saule Suleimenova, [3,4]Nikolai Friesen and [5]Dariga Batayeva**

*[1]Department of Agronomy, Kazakh National Agrarian University, Almaty, Kazakhstan*
*[2]Republican State-Owned Enterprise on the Right of Economic Management "Institute of Botany and Phytointroduction" of Forestry and Wildlife Committee of the Ministry of Ecology, Geology and Natural Resources of the Republic of Kazakhstan, Almaty, Kazakhstan*
*[3]Botanical Garden of the University of Osnabrück, Osnabrück, Germany*
*[4]Department of Pharmaceutical and Natural Sciences, I.M. Sechenov First Moscow State Medical University, Moscow, Russia*
*[5]Department of Biology, Kazakh National Women's Teacher Training University, Almaty, Kazakhstan*

*Article history*
Received: 25-07-2022
Revised: 29-09-2022
Accepted: 25-11-2022

Corresponding Author:
Akerke Tokenova
Department of Agronomy, Kazakh National Agrarian University, Almaty, Kazakhstan
Email: tokenova.akerke@bk.ru

**Abstract:** The genus *Allium* L comprises over 1000 species with important nutritional, medicinal and horticultural applications. The natural flora of Kazakhstan is rich in *Allium* species, including the wild long-pointed onion, *A. longicuspis* and cultivated *A. sativum* garlic varieties. It has been hypothesized that *A. longicupis* is ancestral to *A. sativum*. We investigated genetic polymorphism in samples of *A. longicupis* and the *A. sativum* cultivars "Niki" and "Merey" using start codon targeted polymorphism analysis. Nuclear ribosomal DNA fragments were purified, amplified and tested with 25 primers, finally selecting 12 primers that identified polymorphisms. DNA samples were sequenced, electrophoresed and cluster analysis performed by the construction of dendrograms and principal component analysis. The karyotypes of the different samples were also compared. The results showed that the samples of *A. longicupis* from natural populations did not differ significantly from those of the cultivated garlic varieties. The results indicate that *A. longicupis* can be considered a feral, rather than an ancestral, form of *A. sativum*. These findings have implications both for the breeding and cultivation of garlic varieties and for the taxonomy of the *Allium* genus.

**Keywords:** Garlic, Long-Pointed Onion, SCoT Method, Karyotype, Population

## Introduction

The natural flora of Kazakhstan is rich in useful plants, with a special place taken by onions (*Allium* L.). The genus *Allium* L. is the largest in the Amaryllidaceae J.St.-Hil family, including over 1000 species (Govaerts *et al*., 2005) and is distributed in the temperate latitudes of both hemispheres. In Kazakhstan, onions grow almost everywhere, from deserts to the mountainous alpine belt (Flora of Kazakhstan, 1958). According to Baitenov (2001a; 2001b) these include 140 species of *Allium* L., of which 45 are endemic while other sources have recognized 120-127 onion species of the *Allium* genus (Tarlachkov *et al*., 2020) in Kazakhstan. Most types of onions have nutritional, decorative and medicinal properties; however, many are still underutilized due to a lack of knowledge of their biological characteristics.

The natural populations of the wild race of *Allium sativum*, the long-pointed onion (*Allium longicuspis*), like cultivated garlic, do not form seeds. Fertile seeds of garlic can only be obtained experimentally (Kamenetsky *et al*., 2004; 2007; Shemesh-Mayer and Kamenetsky Goldstein, 2018). *A. longicuspis* is practically indistinguishable from garlic (*A. sativum*) and is currently considered as a synonym for garlic (Shemesh-Mayer and Kamenetsky Goldstein, 2018; Maaß and Klaas, 1995; POWO, 2020). When the wild plants are cultured, they grow rapidly and increase considerably in size.

In the literature, there are two hypotheses on the relationship between these two taxa. According to the first

hypothesis, *A. longicuspis* is considered to be a direct wild ancestor of the cultivated species *A. sativum* (Maaß and Klaas, 1995; Vvedensky, 1935; Etoh and Ogura, 1984; Pooler and Simon, 1993; Mathew, 1996; Hong, 1999) the second hypothesis suggests that *A. longicuspis* is a wild form of cultivated garlic, as it is always found on roads and in abandoned settlements and has perfect conditions for growth in Central Asia (Kamenetsky *et al*., 2004; Fritsch and Friesen, 2002; Etoh and Simon, 2002). Unfortunately, the real wild ancestor of cultivated garlic has not yet been found. It is assumed that the transition to vegetative reproduction was caused by the selection of cultivated garlic (Shemesh-Mayer and Kamenetsky Goldstein, 2018; Kamenetsky *et al*., 2005; 2002).

Cultivated garlic, *A. sativum*, is widely used as a food, medicinal, bactericidal and insecticidal plant. *A. longicuspis* is a perennial plant that grows up to 110 cm high. The lower half of the stem is covered with leaf sheaths. The leaves are broadly linear, flat and 13-41 cm in length. The bulbs are ovoid and the shells of the bulbs are papyraceous or tunicated. The umbel contains numerous black-violet bulbs, mixed with bractlets, with a drooping cover, 3-4 times longer than the umbel, with a long beak (Fig. 1). The species is present in bushes, along stream banks and in gorges of the lower mountain belt of south-eastern and southern Kazakhstan, specifically, in Zailiyskiy Alatau, the Chu-Ili Mountains, Karatau and Western Tien-Shan (Flora of Kazakhstan, 1958) (Fig. 2).

The study of the genetic diversity of cultivated and wild plants is a priority of modern plant genetics. The assessment of the levels of polymorphism in both the whole genome and individual genes and gene families, including genetic variants determining economically valuable features, such as resistance to abiotic stresses and phytopathogens, is of undoubted scientific and practical interest.

To this end, we performed a molecular genetic analysis of natural populations of *A. longicuspis* from the Chu-Ili Mountains and the north-western continuation of the Zailiyskiy Alatau ridge (Northern Tien-Shan) and compared them with the Kazakh garlic cultivars (*A. sativum*) "Niki and Merey". This is the first study of this kind.

The use of highly informative molecular genetic methods makes it possible to assess the state of the species, its genetic origin and its potential in natural populations. In this regard, the study of *A. longicuspis* in natural populations and their comparison with the cultivated Kazakh garlic varieties “Niki” and “Merey” using DNA fragment analysis Polymerase Chain Reaction (PCR)-based methods are of particular relevance (Kamenetsky *et al*., 2005). DNA markers have been widely used in the molecular genetic testing of plants (Volk *et al*., 2009; Gupta *et al*., 1999; Semagn *et al*., 2006). The two most common uses for DNA markers have been to assess genetic diversity within the germplasm and to map the genetic relationships or Quantitative Trait Loci (QTL) that control agronomical important traits (Winter and Kahl 1995; Collard *et al*., 2005) The use of Random Amplified Polymorphic DNA (RAPD) markers (Farooq and Azam, 2002; Welsh and McClelland 1990), Inter-Simple Sequence Repeat markers (ISSR) (Welsh and McClelland, 1990) and Amplified Fragment Length Polymorphism (AFLP) markers (Blair *et al*., 1999) have been popular as these marker methods can generate a relatively large number of DNA markers per sample and are technically simple. In cultivated plants, RAPD, ISSR, and AFLP markers have been widely used for genetic diversity analysis and QTL mapping (Volk *et al*., 2009; Vos *et al*., 1995; Botha and Venter, 2000; Dziechciarkova *et al*., 2004; Gostimsky *et al*., 2005; Kelly and Miklas 1998; Mueller and Wolfenbarger, 1999). Each marker system has certain advantages and disadvantages compared to the others. In many research laboratories, marker systems are greatly affected by the crop species, technical knowledge, available equipment and research funding. Recently, many new and promising marker methods have appeared. These methods include amplified retrotransposon polymorphism, amplified retrotransposon microsatellite polymorphism (Rao *et al*., 2002), Sequence-Associated Amplified Polymorphism (SRAP) (Kalendar *et al*., 1999) and amplified Target Region Polymorphism (TRAP) (Li and Quiros, 2001). Combined with the rapid growth in genomics research, there is a trend to a shift from random DNA markers to gene-targeted markers such as the SCoT method (Start Codon Targeted DNA) (Hu and Vick, 2003). In SCoT, DNA markers are produced by PCR using primers designed from a short-conserved region flanking the ATG start codon, which is conserved in all genes, with the addition of 15 or more randomly selected nucleotides. This method resembles the RAPD or ISSR amplification reaction, as the same primer is used as both the forward and reverse primer (Welsh and McClelland, 1990; Hu and Vick, 2003), with the only difference being in the primer length. The markers are visualized by standard agarose gel electrophoresis and staining, making this method suitable for many research laboratories with standard equipment. The method is reproducible and can be applied to identify intraspecific and interspecific variations, species, populations, varieties, lines and individuals (Hu and Vick, 2003; Gupta *et al*., 1994; Andersen and Lübberstedt, 2003).

The aim of the study was to assess the genetic diversity and determine the population genetic structure of *A. longicuspis* in natural populations and compare them with the "Niki" and "Merey" garlic cultivars from the Kazakh Research Institute of Potato and Vegetable Growing (KazNIIKO) in Almaty using SCoT molecular markers. For this study, fragments of nuclear ribosomal DNA (Internal Transcribed Spacers, ITS) were also sequenced from seven samples of long-pointed onion and six samples of garlic. The karyotypes of the long-pointed onion and the cultivated garlic varieties were also compared.

**Fig.1:** *Allium longicuspis* regel in the Main Botanical Garden collection

**Fig. 2:** *Allium longicuspis* regel from natural populations in the Chu-Ili Mountains (A and B-first population; C-second population)

## Materials and Methods

The study objects were represented by samples of the wild species *A. longicuspis* from two natural populations in the Chu-Ili Mountains and the Kazakh garlic cultivars "Niki" and "Merey". Route-reconnaissance surveys were conducted and material was collected from the two natural populations of the long-pointed onion. The first population was represented by three sites located on the slopes of the Chu-Ili Mountains with heights ranging from 950 to 1069 m above sea level, 7-10 km west of Kurdai village in the territory of the Kordai district of the Jambyl region. The second population, represented by one site, was located 3 km north of Kurdai village in the territory of the Kordai district of the Jambyl region (Table 1).

The plants from the first population were almost identical in appearance (plant height, leaf length and beak length), although plants at the lower population boundary were distinguished by greater growth, indicating more favorable growing conditions. The second population was located in degraded habitat and both the plant height and cap beak length were approximately half those of the first population.

### *DNA Extraction*

The fresh leaves were dried using silica gel. DNA was extracted using the innu PREP Plant DNA kit (Analyst Jena, Germany) according to the manufacturer's instructions. The working DNA was stored at +4°C. The extracted DNA concentration and quality were determined using a max life personal gene analyzer H100 (Barnaul, Russia), according to the manufacturer's instructions.

### *Amplification and Sequencing*

The sequences of the primers used are shown in Table 2. PCR was performed according to the ITS amplification protocol (Gupta and Rustgi, 2004) with a 20 μL reaction mixture with 2 × HS Taq Mix Red (BioZym, Germany), where the mixture composition was 1 μL of the forward primer (ITS-A) (Blattner, 1999) and the reverse primer (ITS-4) (Bruns *et al*., 1991), 10 μL 2 × HS Taq Mix Red (BioZym, Germany) and 8 μL dis $H_2O$ as described. The amplified products were analyzed by electrophoresis on 1.5% agarose gels stained with ethidium bromide. DNA fragments were visualized under UV light on a Gel I × 20 Lmager apparatus (INTAS Science Imaging, Germany) and documented using a Mitsubishi P93D printer (Mitsubishi Elec. Corp., Japan). The amplification products were sent to the Microsynth SeqLab laboratory (Göttingen, Germany) for nucleotide sequencing.

### *SCoT Methods*

To assess genetic polymorphism, four samples (one sample from each of the *A. longicuspis* populations and one sample each from the "Niki and Merey" cultivars) were investigated with 25 SCoT primers. Twelve of the primers that demonstrated polymorphisms were selected for the analysis (Table 2). PCR was performed with a 20 μL reaction mixture with 2 × HS Taq Mix Red, where the mixture composition was 1 μL of DNA and SCoT primer, 10 μL of Red Mix and 8 μL of dis $H_2O$. The PCR reaction was performed in a Professional Thermocycler Biometra amplifier (Germany), according to the following program: Pre-denaturation-01:30 min at 94ºC, then 36 cycles (00:45 min at +94ºC, 00:45 min at +50ºC, 1:30 min at +72ºC) and the final stage - 6:00 min at +72ºC and 90:00 min at 12ºC. DNA separation was performed in agarose gels with an agarose concentration of 1.5% in an electrophoresis chamber in Tris/Borate/EDTA (TBE) buffer using ethidium bromide. The duration of electrophoresis was 3.5-4 h with an electric field voltage of 85V. DNA visualization was performed using the INTAS Science Imaging system and Intag GDS software. The 100 bp-DNA Leiter EXTENDED marker was used as a DNA standard. The electrophoresis results were analyzed for the presence (1) or absence (0) of bands in the gel, followed by matrix development.

### *Comparison of the Karyotypes of Long-Pointed Onion and Garlic*

Roots of *A. longicuspis* and the garlic "Niki" and "Merey" cultivars were used as the karyological research material. The *A. longicuspis* karyotype was studied in the second population samples and the *A. sativum* karyotype in the "Merey" variety.

**Table 1:** Characteristics of analyzed populations of *Allium longicuspis* and garlic cultivars "Niki", "Merey"

| No. | Type, variety | Sampling location and date | Height above sea level, m | Geographical coordinates | Population code |
|---|---|---|---|---|---|
| 1(1-3) | *Allium longicuspis* | Jambyl region, Kordai district, 7-10 km | 950 | 43º17´532"N, | All (1-3) 1-13 |
| | | west of Kurdai village, Kordai Pass, Chu-Ili Mountains. 05/18/2018 | 1049 | 074°51´521"E 43°13´575" N, 074°54´589"E | |
| 2 (4) | *A. Longicuspis* | Jambyl region, Kordai district, 3 km north of Kurdai village, Kordai Pass, Chu-Ili Mountains. 08/29/2019 | 1070 | 43º20'08.3" N, 074º55'42.1" E | All (4) 1-20 |
| 3 | *A. sativum* Niki' | Kazakh Research Institute of Potato and Vegetable Growing (KazNIIKO) 09/28/2018 | – | – | AsN 1-5 |
| 4 | *A. sativum* 'Merey' | Kazakh Research Institute of Potato and Vegetable Growing (KazNIIKO) 09/28/2018 | – | – | AsM 1-5 |

**Table 2:** SCoT primers used and the numbers of markers obtained

| SCoT primer | Sequence (5′-3′) | Number of markers | Monomorphic |
|---|---|---|---|
| SCoT 2 | CAACAATGGCTACCACCC | 10 | 2 |
| SCoT 7 | CAACAATGGCTACCACGG | 11 | 1 |
| SCoT 11 | AAGCAATGGCTACCACCA | 7 | 2 |
| SCoT 12 | ACGACATGGCGACCAACG | 6 | 1 |
| SCoT 17 | ACCATGGCTACCACCGAG | 11 | 1 |
| SCoT 19 | ACCATGGCTACCACCGGC | 14 | 2 |
| SCoT 21 | ACGACATGGCGACCCACA | 5 | 1 |
| SCoT 23 | CACCATGGCTACCACCAG | 15 | 3 |
| SCoT 24 | CACCATGGCTACCACCAT | 8 | 4 |
| SCoT 27 | ACCATGGCTACCACCGTG | 14 | 1 |
| SCoT 30 | CCATGGCTACCACCGGCG | 15 | 2 |
| SCoT 35 | CATGGCTACCACCGGCCC | 15 | 1 |
| Total number of obtained markers | | 131 | 21 |

Chromosomes were studied in the root tip meristem fixed with acetic alcohol (3:1 acetic acid: Ethanol) with preliminary treatment with 8-hydroxyquinoline saturated solution (4 h), followed by washing and storage in 70% ethanol.

The set of somatic chromosomes was determined on temporary squash preparations stained with acetohematoxylin according to Yu. A. Smirnova's method (Govaerts *et al*., 2005). For this, the fixed roots were put in a 4% solution of iron ammonium quartz for 2-3 min. Then, they were dipped into a bottle with the dye and heated to boiling using an alcohol lamp. After the solution had cooled, it was possible to prepare the squash preparation for analysis. The chromosomes were measured on photographs taken with a Nikon D70 camera on a Leitz Aristoplan microscope. For each sample, the average karyotype value was determined based on measurements of at least 10 photographs of metaphase plates. Well-spread metaphase plates were electronically documented (digitally photographed) and, finally, the chromosomes on the best plates were measured and pairwise arranged using the KaryoType software (Altınordu *et al*., 2016). Because the idiograms automatically assembled by the software were not satisfactory, we manually ordered the chromosome pairs according to their size and the position of their centromere. The idiograms were designed using the bar graph function implemented in MS Excel®. For morphological analysis of chromosomes, the centromeric index of each Chromosome (Ci) was calculated using the formula Ci = S: S + L × 100, where S is the absolute length of the short chromosome arm and L is the absolute length of the long chromosome arm. To characterize the karyotype as a whole, the karyotype centromerix index was determined using the formula (KCi = ΣS: ΣS + L × 100. Chromosome sets were divided into types according to (Levan *et al*., 1964) with the modifications described by Vavilov (1966).

## Results

### *Sequences of Ribosomal DNA Fragments*

ITS sequences were obtained from seven *A. longicuspis* (three samples of each population and one sample from Kyrgyzstan from the collection of onions of the Botanical Garden of the University of Osnabrück) and six *A. sativum* samples (two samples of each cultivar and one sample from Germany and Israel from the onion collection of the Botanical Garden of the University of Osnabrück. All sequences were identical to one another and with previously published sequences, as shown by BLAST sequence alignment (www.ncbi.nlm.nih.gov/BLAST). All 13 sequences are published in the NCBI nucleotide sequence database. The samples with their GenBank accession numbers are shown in Table 3.

**Table 3:** Sequenced ITS ribosomal DNA samples and their accession numbers in the GenBank database

| No. | Sample name | GenBank accession number |
|---|---|---|
| 1 | *Allium sativum*_Niki_1 | MT023515.0 |
| 2 | *Allium sativum*_Niki_2 | MT023516.0 |
| 3 | *Allium sativum*_Merey_1 | MT023517.0 |
| 4 | *Allium sativum*_Merey_2 | MT023518.0 |
| 5 | *Allium longicuspis*1-3_1 | MT023519.0 |
| 6 | *Allium longicuspis*1-3_3 | MT023520.0 |
| 7 | *Allium longicuspis*1-3_9 | MT023521.0 |
| 8 | *Allium longicuspis*4_8 | MT023522.0 |
| 9 | *Allium longicuspis*4_12 | MT023523.0 |
| 10 | *Allium longicuspis*4_17 | MT023524.0 |
| 11 | *Allium longicuspis*_Al11_Kirgisia | MT023525.0 |
| 12 | *Allium sativum*_Bamberger_Garlic | MT023526.0 |
| 13 | *Allium sativum*_Israel | MT023527.0 |

*A. sativum* samples 1 and 2 were from the "Niki" cultivar and samples 3 and 4 from the "Merey" cultivar. *A. longicuspis* samples 5-7 were from Population 1 and samples 8-10 were from Population 2. Sample 11 was from the University of Osnabrück, Kyrgyzstan. *A. sativum* samples 12 and 13 were obtained from Germany and Israel, respectively
All 13 sequences were found to be 100% identical

**Table 4:** Morphometric characteristics of the *Allium longicuspis* karyotype from the second population based on 20 metaphase plates

| Pair n. | L | S | Sat | Summa | Ci | Type |
|---|---|---|---|---|---|---|
| I | 6.16 | 5.23 | | 11.39 | 45.90 | m |
| II | 5.56 | 4.78 | | 10.34 | 46.23 | m |
| III | 5.61 | 1.48 | 3.21 | 10.30 | 14.36 | st |
| IV | 5.39 | 4.32 | | 9.700 | 44.53 | m |
| V | 5.10 | 3.75 | | 8.900 | 42.13 | m |
| VI | 4.44 | 1.24 | 2.80 | 8.480 | 21.80 | st |
| VII | 4.76 | 3.44 | | 8.200 | 41.95 | m |
| VIII | 4.33 | 2.75 | | 7.080 | 38.84 | m |
| TKL=148,78 μm; $CI_{2n}$ = 39,47%, $K_{2n}$ = 16 = 12 m +4st | | | | | | |

Pair n.-number of chromosome pair; L-length of chromosome long arm; S-length of chromosome short arm; Sat-satellite; Ci-centromeric index; Type-chromosome type: M-metacentric; st-subtelocentric; TKL-diploid chromosome set length; $CI_{2n}$-centromeric karyotype index

**Table 5:** Morphometric characteristics of the *Allium sativum* variety "Merey" karyotype based on 10 metaphase plates

| Pair n. | L | S | Sat | Summa | Ci | Type |
|---|---|---|---|---|---|---|
| I | 6.72 | 4.30 | | 11.02 | 39.00 | m |
| II | 5.53 | 4.71 | | 10.24 | 45.99 | m |
| III | 5.28 | 4.28 | | 9.560 | 47.76 | m |
| IV | 5.04 | 4.51 | | 9.550 | 47.00 | m |
| V | 5.33 | 1.61 | 2.39 | 9.330 | 23.19 | st |
| VI | 4.58 | 3.58 | | 8.160 | 43.87 | m |
| VII | 4.55 | 0.90 | 2.48 | 7.930 | 16.51 | st |
| VIII | 3.84 | 2.33 | | 6.170 | 37.76 | m |
| TKL=143,86 μm; $CI_{2n}$ = 39,09%, $K_{2n}$ = 16 = 12 m +4st | | | | | | |

Pair n.-number of chromosome pair; L-length of chromosome long arm; S-length of chromosome short arm; Sat-satellite; Ci-centromeric index; Type-chromosome type: M-metacentric; st-subtelocentric; TKL-diploid chromosome set length; $CI_{2n}$-centromeric karyotype index

## *SCoT Analysis*

The analysis of 33 long-pointed onion and 10 garlic samples by the SCoT method yielded 120 polymorphic markers, representing 91% of the total number of traits and 11 monomorphic marker features. No private markers unique to either of the *A. longicuspis* populations or garlic cultivars were identified. The primers used revealed from 6 to 15 polymorphic fragments per sample (Table 3). The lengths of the DNA fragments ranged from 150 to 1800 base pairs. Nevertheless, the first population of long-pointed onion from the Chu-Ili Mountains, consisting of three subpopulations, was found to be more polymorphic in comparison with the second population from the Kordai Pass. This population had a mixed character which is evident in the electrophoretic images (Fig. 3).

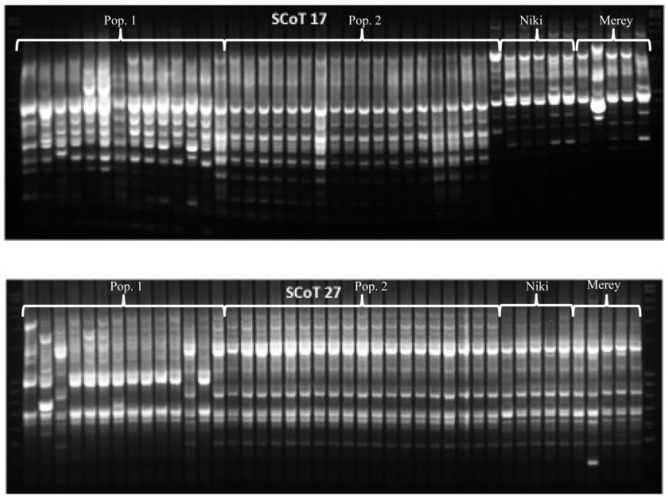



**Fig. 3:** SCoT patterns of the two populations of long-pointed onion and the two garlic varieties. A, SCoT17 primer. B, SCoT27 primer

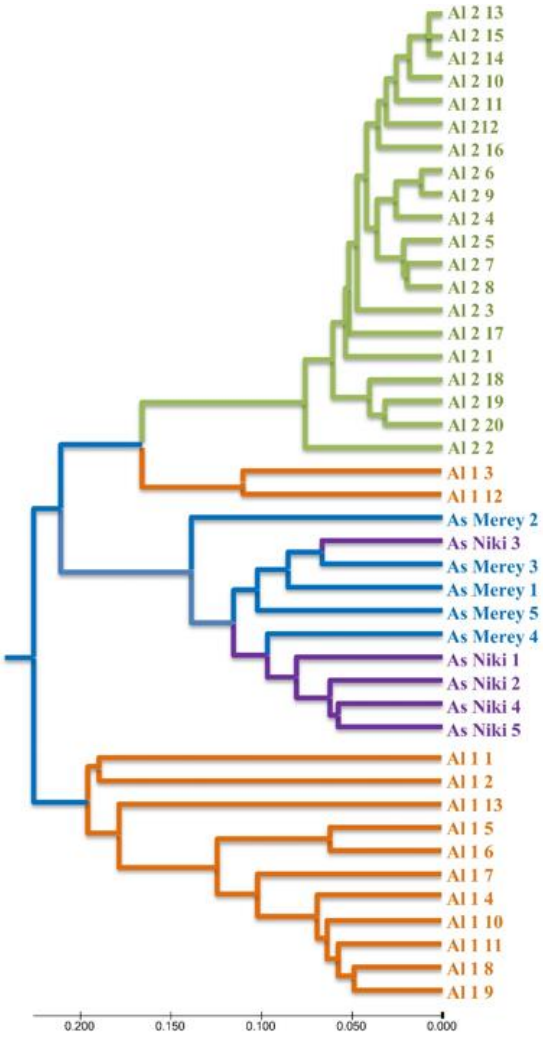



**Fig. 4:** UPGMA dendrogram of the relationships between the two populations of long-pointed onion and the two garlic varieties based on the SCoT markers Al 1 (1-13), *Allium longicuspis*, first population; Al 2 (1-20), *Allium longicuspis*, second population; As-*Allium sativum*

**DOI: 10.3844/ojbsci.2023.33.43**

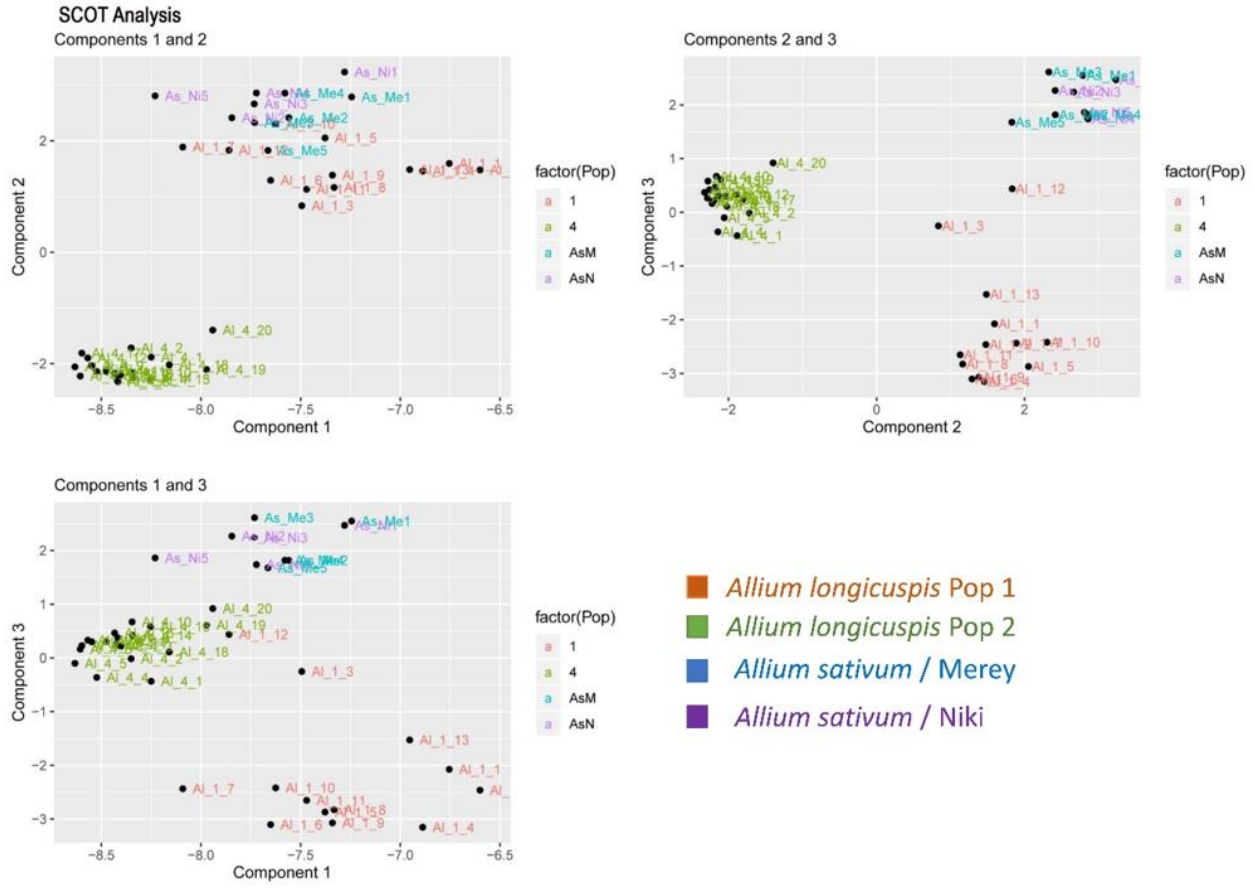
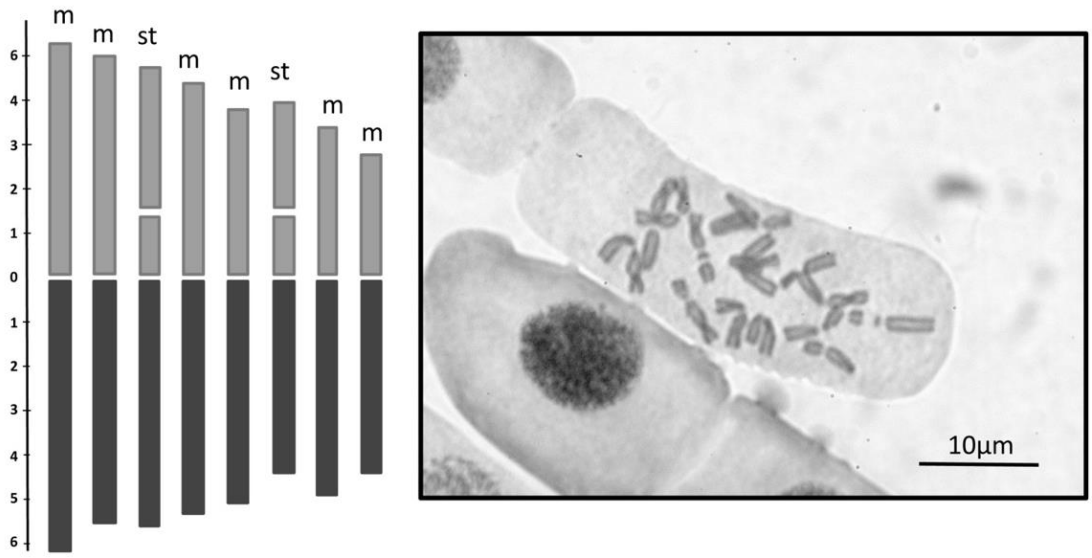
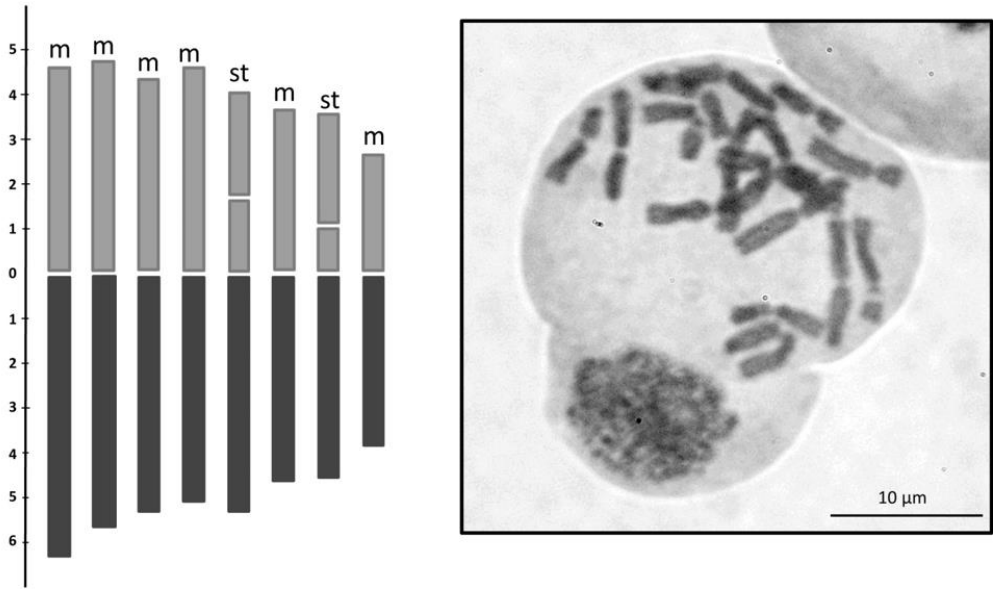



**Fig. 5:** Ordination of analyzed populations based on SCoT markers of the long-pointed onion and two garlic varieties using principal component analysis




(a)

**Fig. 6a:** Morphometric characteristics of the *Allium longicuspis* karyotype from the second population based on 20-metaphase plates




(b)

**Fig. 6b:** Morphometric characteristics of the *Allium sativum* karyotype 'Merey' based on 10-metaphase plates

### *Phylogenetic and Statistical Analysis*

Cluster analysis using dendrograms and PCA based on the SCoT matrix clearly separated the second population from the Kordai Pass from the first population of *A. longicuspis* and the garlic cultivars "Niki and Merey". In the dendrogram (Fig. 4) and on the PCA (Fig. 5), it can be seen that the first *A. longicuspis* population appeared to be more polymorphic, as shown by the longer branches, although not yet monophyletic. Two samples, Al 1-3 and Al 1-12, from the first population appeared to be more closely related to the cultivated garlic varieties and samples from the second population. Both the garlic varieties formed a monophyletic clade but were intermixed. In the dendrogram, the *A. sativum* varieties also appear mixed between the long-pointed onion populations (Fig. 4). The tree based on the degree of similarity of populations by SCoT marker features showed a division into two main branches with one (the lower branch in Fig. 4) including most of the samples from the first long-pointed onion population and the second branch (the upper branch in Fig. 4) including all the other samples of both long-pointed onion and cultivated garlic. This second branch, in turn, is divided into two subgroups: samples from population 2 from the Kordai Pass and the cultivated garlic samples, with two samples from the first *A. longicupis* population, Al_1_3 and Al_1_12, showing an intermediate position between these two subgroups on the dendrogram (Fig. 4). This intermediate position of samples Al_1_3 and Al_1_12 can also be clearly seen in the PCA when comparing the second and third components and the first and third components (Fig. 5). It is interesting that the comparison of components one and two (Fig. 5) shows only two clouds, in which all the samples from *A. longicuspis* population 1 are combined and both populations of *A. sativum* and the *A. longicuspis* population 2 samples are well separated.

The AMOVA results showed identical percentages of molecular variance within and among populations (50 %). The AMOVA results are shown in Supplement 1.

### *Comparison of Karyotypes*

The karyotype of the second long-pointed onion population was developed based on images of 20 metaphase plates (Table 4, Fig. 6a) and the karyotype of the garlic "Merey" variety was based on 10 images of metaphase plates (Table 5 and Fig. 6b). The morphometric characteristics of the analyzed karyotypes do not differ essentially from each other; there are only insignificant discrepancies in the lengths and, accordingly, in the arrangement of satellite chromosome pairs in karyotypes, specifically, the 3$^{rd}$ and 6$^{th}$ pairs for the long-pointed onion and the 5$^{th}$ and 7$^{th}$ pairs for the garlic (Fig. 6). The karyotype formula ($K_{2n}$) and the centromeric index of the chromosome sets ($CI_{2n}$) in the analyzed samples did not differ significantly between *A. longicuspis* and *A. sativum* (Tables 4 and 5).

## Discussion

Garlic is cultivated worldwide and is much in demand for its culinary, nutritional and medicinal properties. Kazakhstan has a rich natural flora that includes many representatives of the *Allium* genus, including both wild and cultivated garlic

and, therefore, presents an ideal environment for investigating the interrelationships between these different varieties. This is an important issue as it affects garlic cultivation and horticulture, particularly in the selection of desirable and economically valuable traits. It has been suggested, on the one hand, that the wild long-pointed onion *A. longicupis* is ancestral to *A. sativum*, the cultivated garlic (Shemesh-Mayer and Kamenetsky Goldstein, 2018; POWO, 2020; Vvedensky, 1935; Etoh and Ogura, 1984; Pooler and Simon, 1993; Mathew, 1996), while, on the other hand, it is possible that *A. longicupis* is a wild form of cultivated garlic (Vavilov, 1966; Hong, 1999; Fritsch and Friesen, 2002; Etoh and Simon, 2002). The use of molecular genetic techniques offers a solution to this question.

These primers were found to be highly effective for the study of *A. longicuspis* and *A. sativum*. Due to their large number and random distribution, SCoT markers give good genetic distance results between the taxa being compared, therefore, for initial screening and structuring of a large collection, SCoT marker analysis is a valuable tool for determining the population structure and related relationships within a species.

Our results showed that the nuclear ribosomal DNA fragment sequences from the four samples (the two *A. longicupis* populations and the two *A. sativum* cultivars) were not only identical to one another, but also to an *A. longicupis* sample from Kyrgyzstan and to two *A. sativum* samples from Germany and Israel, respectively. This is an indication of close genetic relationships between the wild and cultivated species. The SCoT analysis and electrophoresis (Fig. 3) indicated that the first *A. longicupis* population from the Chu-Ili Mountains was more polymorphic than the second population from the Kordai Pass. This polymorphism was confirmed in the dendrogram (Fig. 4).

## Conclusion

The dendrogram compiled by the UPGMA method showed a division of the samples into two main branches. One branch contained samples from the first *A. longicupis* population and the other samples from the second population together with the *A. sativum* cultivars and two samples from the first population. This separation was confirmed by the PCA analysis. The UPGMA method produces rooted trees and the dendrogram (Fig. 4) suggests that, while it is possible that *A. longicupis* may indeed be partly ancestral to cultivated garlic, there is considerable intermixing. However, because we only compared local varieties of *A. sativum* and *A. longicuspis*, we cannot reach a definite conclusion. It is also interesting that, in the second *A. longicupis* population, there appears to be considerable rapid diversification. This population is located in the Kordai Pass, a harsher and more arid environment than that of the first population and it is possible that this diversification may represent adaptation to this environment and even selection by the harsher conditions.

In conclusion, our results confirm the findings of previous researchers (Shemesh-Mayer and Kamenetsky Goldstein, 2018; Etoh and Ogura, 1984; Hong, 1999) that *A. longicuspis* is no different from cultivated garlic and should be considered as feral *A. sativum*. Since the populations of feral garlic demonstrate distinctive molecular signatures from studied garlic varieties, they are a natural reservoir of the genetic diversity of old local garlic forms that have survived in nature. This has potential relevance for garlic cultivation and horticulture. Further molecular genetic study of all long-pointed onion populations is required for further applications in breeding.

## Funding Information

The article was written according to the program: BR10264557 "Cadastral assessment of the current ecological state of the flora and plant resources of the Almaty region as a scientific basis for the effective management of resource potential" for 2022.

## Author's Contributions

**Akerke Tokenova and Gulnar Sitpayeva:** Participated in all experiments coordinated the data-analysis and contributed to the writing of the manuscript.

**Nadezhda Gemejiyeva and Saule Suleimenova:** Designed the research plan and organized the study.

**Nikolai Friesen:** Participated in all experiments.

**Dariga Batayeva:** Contributed to the writing of the manuscript.

## Ethics

This article is original and contains unpublished material. The corresponding author confirms that all of the other authors have read and approved the manuscript and no ethical issues involved.

## References

Altınordu, F., Peruzzi, L., Yu, Y., & He, X. (2016). A tool for the analysis of chromosomes: KaryoType. *Taxon*, 65(3), 586-592.https://doi.org/10.12705/653.9

Andersen, J. R., & Lübberstedt, T. (2003). Functional markers in plants. *Trends in plant science*, 8(11), 554-560. https://doi.org/10.1016/j.tplants.2003.09.010

Baitenov, M. S. (2001a). Flora of Kazakhstan: Generic complex of flora.

Baitenov, M. S. (2001b) Geografo-geneticheskiy analiz lukov (Allium L.) Dzhungarskogo Alatau (V kn.: Izuchenie rastitel'nogo mira Kazakhstana i ego okhrana) [Geographic and genetic analysis of onions (*Allium L.*) of the Dzungarian Alatau//In the book: Study of the plant world of Kazakhstan and its protection.] Mat. Bot.Conf. Almaty.- pp. 19-25. [In Russian] (Baitenov MS Geographic and genetic analysis of onions (*Allium L.*) of the Dzungarian Alatau//In the book: Study of the flora of Kazakhstan and its protection. Mat. Bot. Conf. Almaty. 2001b. 19-25 p.).

Blair, M. W., Panaud, O., & McCouch, S. R. (1999). Inter-simple sequence repeat (ISSR) amplification for analysis of microsatellite motif frequency and fingerprinting in rice (Oryza sativa L.). *Theoretical and Applied Genetics*, 98(5), 780-792. https://doi.org/10.1007/s001220051135

Blattner, F. R. (1999). Direct amplification of the entire ITS region from poorly preserved plant material using recombinant PCR. *BioTechniques*, 27(6), 1180-1186. https://doi.org/10.2144/99276st0

Botha, A. M., & Venter, E. (2000). Molecular marker technology linked to pest and pathogen resistance in wheat breeding. *South African Journal of Science*, 96(5).

Bruns, T. D., White, T. J., & Taylor, J. W. (1991). Fungal molecular systematics. *Annual Review of Ecology and Systematics*, 525-564. https://doi.org/10.1007/BF00224530.

Collard, B. C., Jahufer, M. Z. Z., Brouwer, J. B., & Pang, E. C. K. (2005). An introduction to markers, Quantitative Trait Loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. *Euphytica*, 142(1), 169-196. https://doi.org/10.1007/s10681-005-1681-5

Dziechciarkova, M., Lebeda, A., Dolezalova, I., & Astley, D. (2004). Characterization of *Lactuca spp*. germplasm by protein and molecular markers-a review. *Plant Soil and Environment*, 50(2), 47-58. https://doi.org/10.17221/3680-PSE

Etoh, T., & Ogura, H. (1984). Comparison between *Allium longicuspis* Regel and a fertile garlic clone and their hybrid seeds. In 1984 Abstr Jpn Soc Hort Sci Spring Meeting (pp. 170-171).

Etoh, T., & Simon, P. W. (2002). Diversity, fertility and seed production of Garlic In: Rabinowitch, HD and *L. Allium* Crop science: *Recent advances, CAB International*. https://doi.org/10.1079/9780851995106.0101

Farooq, S., & Azam, F. (2002). Molecular markers in plant breeding-II. Some pre-requisites for use. *Pakistan Journal of Biological Sciences*, 5(10), 1141-1147. https://doi.org/10.3923/pjbs.2002.1141.1147

Flora of Kazakhstan (1958) [Флора Казахстана.] In: Pavlov VN (ed) Alma-Ata: Science, 1958. - t. 2, pp. 290. [In Russian] (Флора Казахстана. Гл. Ред. Павлов Н. В. – Алма – Ата: Наука, 1958. - т.2. 290 с.)

Fritsch, R. M., & Friesen, N. (2002). Evolution, domestication and taxonomy. *Allium Crop Science: Recent Advances*, 5-30. https://doi.org/10.1079/9780851995106.0005

Gostimsky, S. A., Kokaeva, Z. G., & Konovalov, F. A. (2005). Studying plant genome variation using molecular markers. *Russian Journal of Genetics*, 41(4), 378-388. https://doi.org/10.1007/s11177-005-0101-1

Govaerts, R., Kington, S., Friesen, N., Fritsch, R., Snijman, D. A., Marcucci, R., Silverstone-Sopkin, P. A., & Brullo, S. (2005) World checklist of Amaryllidaceae. https://powo.science.kew.org/

Gupta, M., Chyi, Y. S., Romero-Severson, J., & Owen, J. L. (1994). Amplification of DNA markers from evolutionarily diverse genomes using single primers of simple-sequence repeats. *Theoretical and applied Genetics*, 89(7), 998-1006. https://doi.org/10.1007/BF00224530.

Gupta, P. K., & Rustgi, S. (2004). Molecular markers from the transcribed/expressed region of the genome in higher plants. *Functional & Integrative Genomics*, 4(3), 139-162. https://doi.org/10.1007/s10142-004-0107-0

Gupta, P. K., Varshney, R. K., Sharma, P. C., & Ramesh, B. (1999). Molecular markers and their applications in wheat breeding. Plant breeding, 118(5), 369-390. https://doi.org/10.1046/j.1439-0523.1999.00401.x

Hong, C. J. (1999). Fundamental studies on crossbreeding in garlic, *Allium sativum L*. Kagoschima University, Kagoshima, Japan.

Hu, J., & Vick, B. A. (2003). Target region amplification polymorphism: A novel marker technique for plant genotyping. *Plant Molecular Biology Reporter*, 21(3), 289-294. https://doi.org/10.1007/BF02772804

Kalendar, R., Grob, T., Regina, M., Suoniemi, A., & Schulman, A. (1999). IRAP and REMAP: Two new retrotransposon-based DNA fingerprinting techniques. *Theoretical and Applied Genetics*, 98(5), 704-711. https://doi.org/10.1007/s001220051124

Kamenetsky, R., Khassanov, F., Rabinowitch, H. D., Auger, J., & Kik, C. (2007). Garlic biodiversity and genetic resources. *Medicinal and Aromatic Plant Science and Biotechnology*, 1(1), 1-5.

Kamenetsky, R., London Shafir, I., Khassanov, F., Kik, C., Van Heusden, A. W., Vrielink-van Ginkel, M., ... & Rabinowitch, H. D. (2005). Diversity in fertility potential and organo-sulphur compounds among garlics from Central Asia. *Biodiversity & Conservation*, 14(2), 281-295. https://doi.org/10.1007/s10531-004-5050-9

Kamenetsky, R., Shafir, I. L., Baizerman, M., Khassanov, F., Kik, C., & Rabinowitch, H. D. (2002, August). Garlic (*Allium sativum* L.) and its wild relatives from Central Asia: Evaluation for fertility potential. In XXVI International Horticultural Congress: Advances in Vegetable Breeding 637 (pp. 83-91). https://doi.org/10.17660/ActaHortic.2004.637.9

Kamenetsky, R., Shafir, I. L., Zemah, H., Barzilay, A., & Rabinowitch, H. D. (2004). Environmental control of garlic growth and florogenesis. *Journal of the American Society for Horticultural Science*, 129(2), 144-151. https://doi.org/10.21273/JASHS.129.2.0144

Kelly, J. D., & Miklas, P. N. (1998). The role of RAPD markers in breeding for disease resistance in common bean. *Molecular Breeding*, 4(1), 1-11. https://doi.org/10.1023/A:1009612002144

Levan, A., Fredga, K., & Sandberg, A. A. (1964). Nomenclature for centromeric position on chromosomes. *Hereditas*, 52(2), 201-220. https://doi.org/10.12705/653.9

Li, G., & Quiros, C. F. (2001). Sequence-related amplified polymorphism (SRAP), a new marker system based on a simple PCR reaction: Its application to mapping and gene tagging in Brassica. *Theoretical and Applied Genetics*, 103(2), 455-461. https://doi.org/10.1007/s001220100570

Maaß, H. I., & Klaas, M. (1995). Infraspecific differentiation of garlic (*Allium sativum L.*) by isozyme and RAPD markers. *Theoretical and Applied Genetics*, 91(1), 89-97. https://doi.org/10.1007/BF00220863

Mathew, B. (1996). A review of Allium section Allium. Royal Botanic Gardens, Kew.

Mueller, U. G., & Wolfenbarger, L. L. (1999). AFLP genotyping and fingerprinting. *Trends in Ecology & Evolution*, 14(10), 389-394. https://doi.org/10.1016/S0169-5347(99)01659-6

Pooler, M. R., & Simon, P. W. (1993). Characterization and classification of isozyme and morphological variation in a diverse collection of garlic clones. *Euphytica*, 68(1), 121-130. https://doi.org/10.1007/BF00024161

POWO. (2020) Plants of the World online. *Allium longicuspis* Regel. http://www.plantsoftheworldonline.org/taxon/urn:lsid:ipni.org:names:528333-1

Rao, K. K., Lakshminarasu, M., & Jena, K. K. (2002). DNA markers and marker-assisted breeding for durable resistance to bacterial blight disease in rice. *Biotechnology Advances*, 20(1), 33-47. https://doi.org/10.1016/S0734-9750(02)00002-2

Semagn, K., Bjørnstad, Å., & Ndjiondjop, M. N. (2006). An overview of molecular marker methods for plants. *African Journal of Biotechnology*, 5(25).

Shemesh-Mayer, E., & Kamenetsky Goldstein, R. (2018). Recent advances in sexual propagation and breeding of garlic. Hort Rev, 46, 1-38. https://doi.org/10.1002/9781119521082.ch1

Tarlachkov, S. V., Epiktetov, D. O., Sviridov, A. V., Shushkova, T. V., Ermakova, I. T., & Leontievsky, A. A. (2020). Draft Genome Sequence of Glyphosate-Degrading Achromobacter insolitus Strain Kg 19 (VKM B-3295), Isolated from Agricultural Soil. *Microbiology Resource Announcements*, 9(17), e00284-20. https://www.plantarium.ru/page/floras/of/145.html)

Vavilov, N. I. (1966). Izbrannye sochineniya: Genetika i sele ktsiya. Moskva: Kolos.

Volk, G. M., Richards, C. M., Henk, A. D., Reilley, A., Miller, D. D., & Forsline, P. L. (2009). Novel diversity identified in a wild apple population from the Kyrgyz Republic. *HortScience*, 44(2), 516-518. https://doi.org/10.21273/HORTSCI.44.2.516

Vos, P., Hogers, R., Bleeker, M., Reijans, M., Lee, T. V. D., Hornes, M., ... & Zabeau, M. (1995). AFLP: A new technique for DNA fingerprinting. *Nucleic Acids Research*, 23(21), 4407-4414. https://doi.org/10.1093/nar/23.21.4407

Vvedensky, A. I. (1935) Rod Allium L. vo Flore SSSR [ Genus Allium L. In Flora of USSR.] Leningrad, 4. 112 -280. [In Russian] Введенкий А. И. Род Allium L. во Флоре СССР. Ленинград, Изд-во АН СССР. 1935. т. 4. с. 112-280

Welsh, J., & McClelland, M. (1990). Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Research*, 18(24), 7213-7218. https://doi.org/10.1093/nar/18.24.7213

Winter, P., & Kahl, G. (1995). Molecular marker technologies for plant improvement. *World Journal of Microbiology and Biotechnology*, 11(4), 438-448. https://doi.org/10.1007/BF00364619