Brazilian Journal of Biology®

ISSN 1519-6984 (Print)
ISSN 1678-4375 (Online)

Original Article

Population analysis of psammophyte species *Allium sabulosum* Steven ex Bunge in Kazakhstan

Análise populacional de espécies psamófitas *Allium sabulosum* Steven ex Bunge no Cazaquistão

D. Sh. Abdildanov[a,b*], P. V. Vesselova[a], G. M. Kudabayeva[a], V. F. Zaikov[c], B. B. Osmonali[a], M. V. Skaptsov[c], S. Ussen[a,b], T. S. Kulboldin[a] and N. Friesen[d]

[a]Institute of Botany and Phytointroduction, Almaty, Kazakhstan
[b]Al-Farabi Kazakh National University, Faculty of Biology and Biotechnology, Almaty, Kazakhstan
[c]Altai State University, Faculty of Biology, South Siberian Botanical Garden, Barnaul, Russia
[d]Osnabruck University, Botanical Garden, School of Biology/Chemistry, Osnabruck, Germany

Abstract

The article presents the results of interpopulation analysis, phylogeny and ploidy study of *Allium sabulosum* Steven ex Bunge within the deserts of Kazakhstan. The studied taxon is the type species of the section *Eremoprasum* (Kamelin) F. O. Khassanov, R. M. Fritsch et N. Friesen and belongs to the subgenus *Allium* L, the main range of which covers the Iranian-Turanian and Central Asian desert regions. The studied samples from 10 populations were collected in 5 floristic regions of Kazakhstan, covering the following regions of the Republic: Mangistau, Kyzylorda, Karaganda, Ulytau, Zhambyl and Almaty. To determine ploidy, the collected samples were studied by flow cytometry. The results of nrITS and chloroplast fragment sequencing were used to construct a phylogenetic tree, which also includes data from the NCBI database. Of the 10 populations studied, samples from population 3 from the Balkhash-Alakol floristic region were particularly different. In this population, a hybrid (the second parent, which is unknown and was not included in the analysis) and samples of different ploidy (polyploid) were found. Tetraploid samples were found in the territory of the Betpak-Dala floristic region.

Keywords: *Allium sabulosum,* hybrid, irano-turanian desert, population, section *Eremoprason*.

Resumo

O artigo apresenta os resultados da análise interpopulacional, estudo filogenético e de ploidia de *Allium sabulosum* Steven ex Bunge nos desertos do Cazaquistão. O táxon estudado é a espécie-tipo da secção *Eremoprasum* (Kamelin) F. O. Khassanov, R. M. Fritsch et N. Friesen e pertence ao subgênero *Allium* L, cuja área de distribuição principal abrange as regiões desérticas do Irã-Turânia e da Ásia Central. As amostras estudadas de 10 populações foram colhidas em 5 regiões florísticas do Cazaquistão, abrangendo as seguintes regiões da República: Mangistau, Kyzylorda, Karaganda, Ulytau, Zhambyl e Almaty. Para determinar a ploidia, as amostras recolhidas foram estudadas por citometria de fluxo. Os resultados do sequenciamento de nrITS e de fragmentos de cloroplastos foram utilizados para construir uma árvore filogenética, que também inclui dados da base de dados NCBI. Das 10 populações estudadas, as amostras da população 3 da região florística de Balkhash-Alakol eram particularmente diferentes. Nesta população, foi encontrado um híbrido (o segundo progenitor, que é desconhecido e não foi incluído na análise) e amostras com diferentes níveis de ploidia (poliploide). Foram encontradas amostras tetraploides no território da região florística de Betpak-Dala.

Palavras-chave: *Allium sabulosum*, híbrido, deserto iraniano-turaniano, população, secção *Eremoprason*.

1. Introduction

The genus *Allium* L. (Amaryllidaceae J.St.-Hil.) includes about 1000 species belonging to 15 subgenera and more than 90 sections (The Royal Botanic Gardens Kew, 2025; Friesen, 2023) and belongs to the "petaloid lily monocots" (APG, IV, Chase et al., 2016). The largest centers of diversity of species of this genus are located in Central Asia, as well as in the Mediterranean basin (Fritsch and Friesen, 2002; Choi and Oh, 2011; Khassanov, 2018; Friesen et al., 2024).

In "Flora of Kazakhstan" (Pavlov and Poljakov, 1958) Pavlov N.V. and Polyakov P.P. listed 108 species of onions for the territory of the republic, 29 of which are endemic to it. In later sources, from 120 (Abdulina, 1999) to 140 species (Baitenov, 2001) of the genus *Allium* are indicated for the flora of Kazakhstan. According to the latest data, the genus is represented by at least 130 species on the territory of Kazakhstan (Epiktetov, 2025).

*e-mail: abdildanov00@mail.ru
Received: June 25, 2025 – Accepted: September 17, 2025
Editor: Jairo Lizandro Schmitt

In the desert zone, the species diversity of the genus is small and amounts to about 30 species. Accordingly, there are very few publications reflecting the results of studies of the genus *Allium* growing in arid conditions, including at the population level. An example of such studies can be the work on the study of the types of individual ontogenesis of the coenopopulation of *Allium sabulosum* Steven ex Bunge, which was represented by its type VI with rejuvenated rudiments (Cheryomushkina, 2006) and the publication on *Allium lehmannianum* Merkl. ex Bunge (Abdildanov et al., 2025).

For the population study of desert species of the genus *Allium*, *Allium sabulosum*, a species with a wide range of growth, was chosen. The studied taxon is the type species of the section *Eremoprasum* (Kamelin) F. O. Khass., R. M. Fritsch et N. Friesen of the subgenus *Allium* (Friesen et al. 2006), which has 18 sections with more than 375 species and 35 subspecies and is the richest in the genus (Khassanov, 2018). It should be noted here that the name of the section *Eremoprasum* (dessert onions) chosen by Kamelin (1973), and where *Allium sabulosum* is the type species of the section, can lead to some confusion with the species with the same name – *Allium eremoprasum* Vved., which most authors attribute to the section *Coerulea* F. O. Khass. (Kamelin, 1973; Khassanov, 2018). In our previous work on *Allium lehmannianum* (Abdildanov et al., 2025), it was shown for the first time that *Allium eremoprasum* is related to *A. sabulosum* and that *A. eremoprasum* has little in common with species from the section *Coerulea*.

One of the first to describe the species under study was Alexander Georg von Bunge (1803-1890), whose samples were collected by Christian Steven in the Lower Volga region (to the east of the Astrakhan region and the Volga River in the Ryn sands) (Sagalayev and Firsov, 2014).

In general, *Allium sabulosum* belongs to the group of psammophytic Turanian elements of flora, covering with its range the territories of the South (sands of the Karakum and southern Kuzylkum) and Northern Turan (from the south of the European part of Russia to the eastern Pribalkhashye, occurring in almost all the deserts of Kazakhstan, including the northern Kyzylkum). Meanwhile, due to the anthropophilic nature of the species, the Ulytau mountain range, located in the steppe zone, penetrates through sand massifs, and in the south - Northern Iran. Within Kazakhstan, *Allium sabulosum* (Pavlov and Poljakov, 1958) is found in the following floristic regions: Caspian, Bukeyevsky, Emba, Turgai, Northern Ust-Urta, Mangyshlak, Priaral, Kyzylorda, Betpak-Dala, Moyynkum, Balkhash-Alakol, Kyzylkum, Turkestan, Chu-Ili Mountains, Karatau. At the same time, it is usually confined to sandy-desert habitats, occurring quite often in interdune depressions. The altitudinal range of the species varies from 80 to 510 m above sea level. Such ecological and geographical preferences of *Allium sabulosum* allow it to be classified as a member of the group of Iranian-Turanian elements of desert flora (Kamelin 1973; Kurochkina 1978). Accordingly, the morphological features of the studied species reflect its adaptation to survival in sandy desert conditions (Figure 1). Thus, the sandy onion has small (about 3 mm long) perianth leaflets, greenish or whitish; the leaves are fistulous, linear, with a tip curled into a ring; the outer shells of the bulbs are leathery, with depressed large veins, due to which the surface of the bulb seems longitudinally corrugated; the seeds have an oval-angular shape (Figure 1G), anticlinal walls of the U-type, and the pericanal walls have more or less dense granules (Yusupov et al., 2022) (Figure 1H).

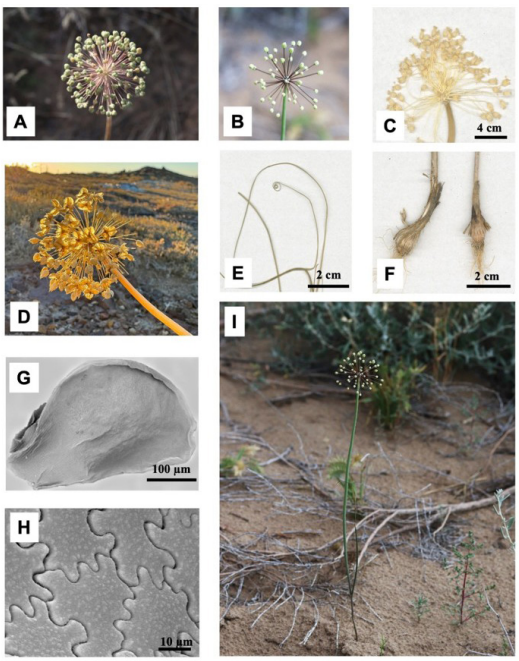

Figure 1. Object of study *Allium sabulosum* (A, B – plants in the budding phase, C – flowering phase, D – fruiting phase, E – leaves, F – bulb, G – plant seeds, H – electron micrograph of the scanned surface of the seed, I – general view of the plant).

Initially, the species was studied using flow cytometry, which is one of the most common methods for studying hybridogenic processes that manifest themselves as polyploidy and aneuploidy. The detection of such hybrid and polyploid individuals in our samples required additional analyses. For more accurate identification of hybrid and polyploid samples, it was decided to use molecular genetic methods to analyze internal transcribed spacers (ITS) and chloroplast fragments. The aim of the present research is to study the population genetic diversity of *Allium sabulosum* in arid regions of Kazakhstan and its position in the classification of the subgenus *Allium*.

2. Materials and Methods

The object of the study is *Allium sabulosum* Steven ex Bunge samples from different populations of the desert regions of Kazakhstan (Figure 1).

Classical botanical methods were used in the study. The following fundamental reference books were used to identify the collected material: "Flora of Kazakhstan" (Pavlov and Poljakov, 1958), "Illustrated Guide to Plants of Kazakhstan" (Illustrated Identifier of Plants of Kazakhstan, 1969, 1972) and "Guide to Plants of Central Asia" (Vvedensky, 1971). The herbarium was collected according to the Skvortsov (1977) method. Population analysis of the species using morphological methods (Mamyrova et al., 2025). The names of plant species were checked against the databases of the International Plant Names Index (IPNI, 2025) and the Plant of the World (POWO, The Royal Botanic Gardens Kew, 2025). The QGIS 3.34 program created a map of *Allium sabulosum* locations. A Jeol JSM 6390LA scanning electron microscope was used to study the surface morphology of the seeds of the studied species. The materials for the study were samples from 10 populations of the *Allium sabulosum* species (Table 1), collected during field studies in the period 2023-2024. The map of population locations is shown in Figure 2A, B.

Within Kazakhstan, *Allium sabulosum* is found in arid and subarid regions, covering: Ulytau, Karaganda, Mangyshlak, Kyzylorda, Zhambyl, Karaganda and Almaty regions of the republic. All these regions, according to modern botanical and geographical zoning, lie within the North Turan province of the Irano-Turan subregion of the Sahara-Gobi desert region (Rachkovskaya et al., 2003).

Table 1. Collection points of samples of the studied *Allium sabulosum* populations.

population №	Floristic areas	Administrative unit (region)	Collection points	Date of collection	N	E	Height (m.a.s.l.)
1	Balkhash-Alakolsky	Almaty	Zhambyl district, Taukuma sands.	03.V.2024	44.362778	75.533056	440
2			Balkhash district, Malaysary mountains, sands.	09.V.2023	44.359167	76.861389	510
3			Balkhash district, left bank of the Ili river, barchans on the right side of the road behind Topar settlement.	28.V.2023	44.946111	75.055556	360
4	Kyzylorda	Kyzylorda	Shiela district, barchans.	20.V.2023	44.273611	66.584167	360
5	Mangyshlak	Mangistau	Mangistau district, near Shetpe village, Northern Aktau, near Shomanai mountain.	14.IV.2024	44.260278	52.105833	80
6	Betpakdala	Karaganda	Aktogay district, northern part of Betpakdala, near Gulshat village, sand.	29.IV.2024	46.673889	74.365833	320
7			Shet district, western part of Betpakdala, sands.	28.IV.2024	46.135000	73.178333	420
8			Shet district, 60 km from Priozersk, western part of Betpakdala, sands.	28.IV.2024	46.070278	72.927778	450
9			Shetsky District, west of Priozersk.	27.IV.2024	46.026667	73.199722	340
10	Moyinkum	Zhambyl	Moyinkum district, near Karl Marx settlement, sands.	24.IV.2024	44.228889	72.787778	350

2.1. Molecular genetics methods

For genetic analysis, three *A. sabulosum* samples were selected from each examined population. DNA extraction was performed using a commercial DiamondDNA kit according to the manufacturer's instructions (OOO Scientific and Production Firm Altaibiotekh, Barnaul).

2.2. Fingerprints analysis iPBS (inter-primer binding sites) retrotransposon markers

To identify genetic diversity, iPBS retrotransposon markers were used (Kalendar et al., 2010). Amplification was performed in 20 µl of the reaction mixture using the BioMaster HS-Taq PCR-Color (2×) PCR kit (OOO Biolabmix, Novosibirsk) in the following composition per sample: 8 µl H2O; 10 µl 2x PCR buffer; 1 µl 10 mM primer; 1 µl total DNA. Amplification was carried out in a BioRad My Cycler thermal cycler using the following protocol: 95 °C (5 min); 30 cycles: 95 °C (20 sec), 57 °C (35 sec), 72 °C (90 sec); 72 °C (7 min). Electrophoresis of the amplification products was carried out in 2% agarose gel, the results were visualized using a transilluminator. Of the 12 primers studied, the five most informative were selected (Table 2).

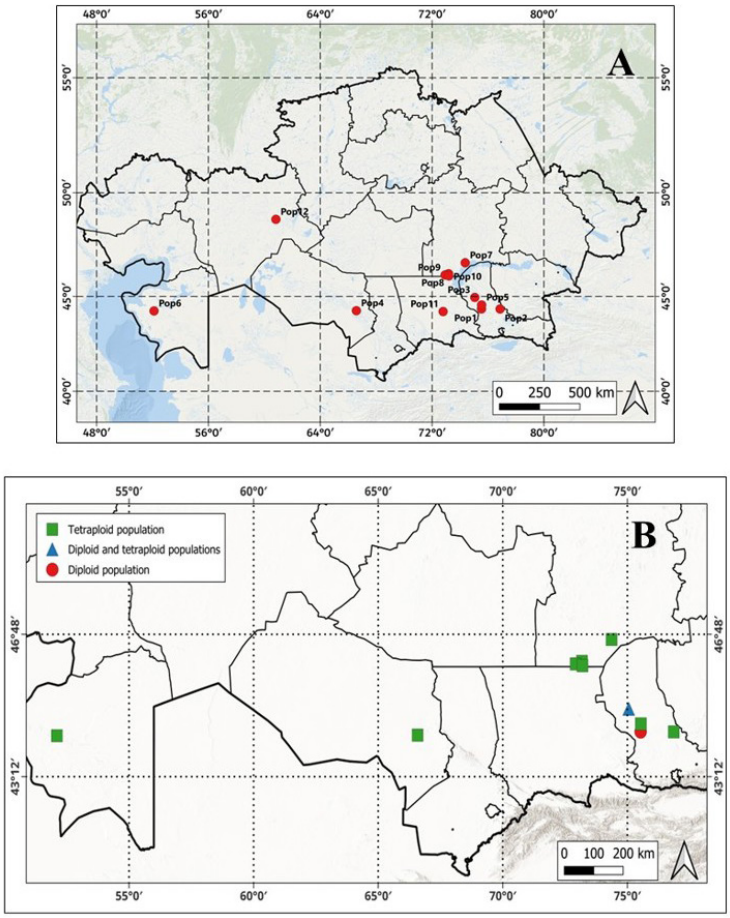

Figure 2. Map of locations of the studied populations of *A. sabulosum*. (**A**-Populations studied, **B**–ploidy distribution).

Table 2. iPBS primers used in the study.

Name of primer	Nucleotide sequence of primer (5'-3')	T (°C) of primer annealing	Number of Amplified DNA fragments	Number of polymorphic DNA fragments (%)
iPBS2239	ACCTAGGCTCGGATGCCA	55	48	83.32
iPBS2272	GGCTCAGATGCCA	50	41	80.5
iPBS2373	ACCTAGCTCATCATGCCA	55	34	82.3
iPBS2237	CCCCTACCTGGCGTGCCA	55	45	84.4
iPBS 2224	ATCCTGGCAATGGAACCA	55	32 (87,5)	

The electropherogram was transformed into a binary data matrix, the presence or absence of a DNA fragment (conditionally a feature) was designated accordingly as 1 or 0. The resulting matrix was analyzed using the TFPGA version 1.3 program (Miller, 1997). The dendrogram was constructed using the unweighted pair-group method (UPGMA – unweighted pair-group method using arithmetic average), the bootstrap coefficient was calculated based on 100,000 iterations. The genetic differentiation matrix was obtained using the GENEPOP program. The genetic structure was analyzed using the STRUCTURE 2.3.4 program (Pritchard et al., 2000) with the following parameters: admixture model with allele frequencies correlated, burn-in 100,000 and run length 100,000 MCMC, the number of clusters (K) from 2 to 12. The optimal number of clusters identified using the STRUCTURE HARVESTER program was 5 (Evanno et al., 2005). The clustering matrix was processed using the ClustVis program (Metsalu and Vilo, 2015). The population structure in pre-defined locations was estimated using two different approaches. First, we applied multivariate methods without relying on a specific population genetics model. Principal component analysis (PCA) was performed in GenAlEx (Peakall and Smouse, 2006). For individual samples, a pairwise genetic distance matrix was calculated, on the basis of which PCA was performed using the covariance standardization method.

Second, population structure analysis was performed using a Bayesian clustering approach implemented in STRUCTURE v2.3.4 (Porras-Hurtado et al., 2013). The analysis was performed within the Admixture model with a burnin period of 20,000 replicates followed by 100,000 Markov Chain Monte Carlo (MCMC) replicates. The number of clusters (K) varied from 2 to 10, with three replicates performed for each K. The data were analyzed in STRUCTURE Harvester (Earl and VonHoldt, 2012) and the optimal K was determined using Evanno et al. (2005) method, which revealed a pronounced peak in the Delta K distribution.

2.3. Cytometry analyses

The DNA content was determined by flow cytometry techniques with propidium iodide (PI) staining. Leaves dried with silica gel were used as samples. Samples were chopped with standard using a sharp razor blade in LB01 buffer containing PI (50 µg/ml), RNase (50 µg/ml) (Doležel et al., 1992) supplemented with 12 mM sodium thiosulfate and 1% polyvinylpyrrolidone (Skaptsov et al., 2024). The nuclear suspension was filtered through nylon filter with a pore size 30 µm. Analyses were performed on a Cytoflex (Beckman Coulter, Inc.) cytometer. Peaks with at least 1000 nuclei and a CV of less than 5% were used for analysis. Histograms were visualized and processed using CytExpert software (Beckman Coulter, Inc.). Descriptive statistic was calculated using XLStat (Addinsoft). As an internal standard was used the Pisum sativum 'Ctirad', 2C = 9.09 pg and Vicia faba 'Inovec', 2C = 26.9 pg (Doležel et al., 1992, 1998).

2.4. Phylogenetic analysis

For one sample from each population, two DNA fragments were sequenced: ITS nuclear and intergenic chloroplast (rpL32-trnL) DNA fragments. For the ITS fragments, primers ITS1 and ITS4 (White et al., 1990; Silbiger et al., 2012) were used. For the intergenic spacer rpL32-trnL, primers rpL32F and trnL UAG (Shaw et al., 2007) were used. Polymerase chain reaction was carried out in 50 µl of the reaction mixture using the Biomaster HS-Taq PCR-Color 2x PCR kit (Biolabmix LLC, Novosibirsk) in the following composition per sample: 25 µl of the ready-made PCR mixture, 21 µl of H2O, 1 µl of 10 mM of the corresponding primers, 2 µl of total DNA. The amplification protocol was: 95°C (3 min); 35 cycles: 95°C (20 s), 57°C (30 s), 72°C (30 s); 72°C (5 min). The amplification products were purified using microcolumns. Sequencing was performed using the Sanger method using an ABI PRISM 3500 XL sequencer. The obtained nucleotide sequences were aligned using the ClustalW algorithm in the MEGAX program (Kumar et al, 2024) with manual read quality assessment. Both data sets (nrITS and the cpDNA (*trn*L-*rpl*32) were analysed separately for position identification in the subgenus *Allium* evolutionary lineage and to find the relationships within subgenus *Allium* through parsimony (PAUP) and Bayesian phylogenetic analysis (MrBayes). Fitch parsimony was determined with the heuristic search option in PAUP version 4.0b10 (Swofford, 2002) with MULTREES, TBR branch swapping and 100 replicates of random addition sequence. Gaps were treated as missing data. The most parsimonious trees returned by the analysis were summarized in one consensus tree using the strict consensus method. Bootstrap analyses using 1,000 pseudoreplicates were performed to assess the support (BS) of the clades (Felsenstein, 1985). Bayesian phylogenetic analyses were also performed using MrBayes 3.1.23 (Ronquist and Huelsenbeck, 2003). The sequence evolution model was chosen following the Akaike Information Criterion (AIC) obtained from jModelTest2 (Darriba et al., 2012). Two independent analyses with four Markov chains were run for 10 million generations, sampling trees every 100 generations. The first 25% of trees were discarded as burn-in. The remaining trees were combined into a single data set, and a majority-rule consensus tree was obtained along with posterior probabilities (PP).

3. Results

3.1. Fingerprints population analysis

During the study of 10 populations of the species *A. sabulosum*, 5 most informative iPBS primers were used. A total of 200 iPBS fragments were identified. The largest number of amplified DNA fragments belonged to the iPBS2239 primer – 48 fragments, the smallest to the iPBS2224 primer – 32 fragments. The indicators of genetic difference (Fst coefficient) between the populations are presented in Appendix A (Supplementary Material).

The population similarity dendrogram constructed on the basis of iPBS fragments in the TFPGA program is shown in Figure 3. The closest were populations 6–9 growing in the Betpak-Dala floristic region (Table 1.), as well as population 10 collected in the Moiynkum floristic region (Table A1). The closest to them were populations 1 and 2 from the Balkhash-Alakol floristic region.

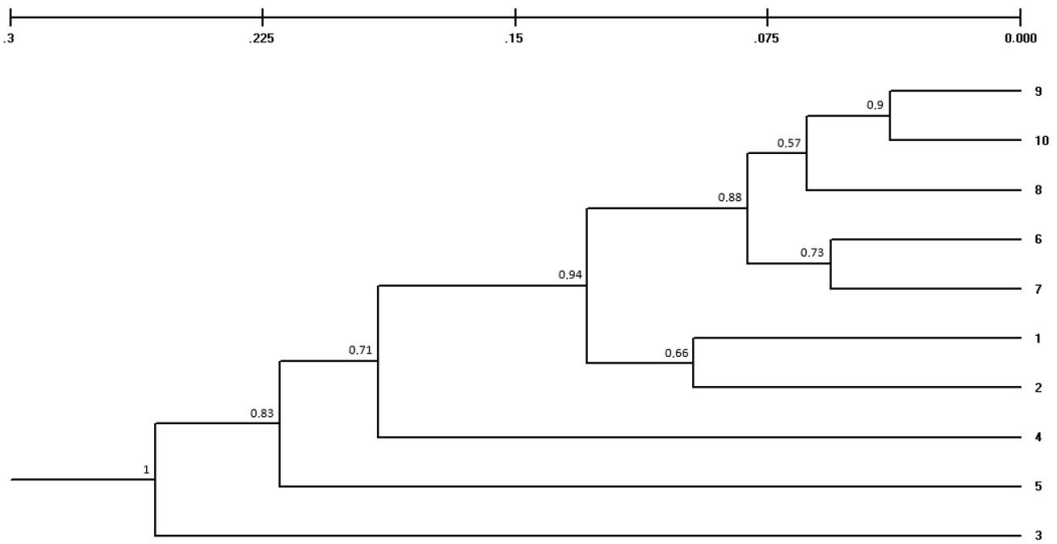

Figure 3. UPGMA Population similarity dendrogram based on iPBS-PCR analysis (nodes indicate bootstrap support value).

Populations 4 (Kyzylorda floristic region) and 5 (Mangyshlak floristic region) stand apart from all the others, which is due to their geographic remoteness from the other studied populations.

Unexpected data were obtained for population 3 from the Balkhash-Alakol floristic region. Two of the three samples of the population were especially different from all the other studied samples. These samples are assumed to be a hybrid, but since the second parent is currently unknown (and therefore was not included in the analysis), the third population was further away than all the others. Hybrid origin is suggested by morphological data: loose inflorescence, unequal peduncles and fibrous bulb coat.

The STRUCTURE analysis determined the most probable number of clusters (K) by calculating the logarithm of the probability of the data for each K value. The highest probability (ProbK = 1.00) was recorded at K = 5 (Appendix B, Figure B1). The Evanno method implemented in STRUCTURE Harvester was used to determine the optimal K value for the genotyped *A. sabulosum* lines. According to the cluster model, the Evanno test showed that at K = 5, the ΔK value was 24.676, which corresponds to the highest logarithmic probability and confirms the presence of five main groups (subpopulations) in the collection (Appendix B, Figure B2).

The population structure was constructed to reveal its architecture. The histogram from STRUCTURE shows the genetic clustering of individuals from ten populations for different values of K (from 2 to 10), reflecting the number of putative genetic clusters (Figure 4).

For K = 2, most populations are predominantly assigned to a single genetic cluster, with the exception of population 3 and, to a lesser extent, population 5, where admixture is observed. As K increases, additional levels of structure are revealed, particularly in populations 3, 4, and 5, which exhibit high levels of genetic heterogeneity. The optimal value of K, previously determined to be K = 5, allows for clear genetic clusters to be identified, although some populations still show evidence of admixture.

Populations 6–10 and 1–2 showed the greatest genetic similarity, consistent with their geographic proximity. In contrast, populations 4 and 5 showed significant genetic divergence, also consistent with their geographic distribution. The samples from population 3 originating from the Balkhash-Alakol floristic region showed genetic heterogeneity: two of the three samples had a unique genetic composition compared to the other analyzed samples. At higher K values, population 3 showed enhanced substructuring, which may be related to its hybrid origin.

The principal component analysis (PCA) plot shows the genetic relationships between the ten populations based on the first two principal components (PC1 and PC2), which explain 21.7% and 14.4% of the total genetic variance, respectively (Figure 5).

3.2. Polymorphism in nrITS sequences in Allium sabulosum

Internal transcribed spacers (nrITS) and chloroplast fragments (trnL-rpL32 spacer) were sequenced from 10 *Allium sabulosum* populations. ITS fragments for 3 populations consisted of two samples, and the others consisted of one sample. The ITS fragment is 643–645 bp long, and the rpL32–trnL fragment is 838–873 bp. Comparison of the sequences revealed some differences in the studied samples.

Thus, in the ITS fragment of the sample from the fifth population, there are individual substitutions of thymine for cytosine at positions 53, 259, 602, 605, and adenine for guanine at position 422. The sample from the fourth population differs from the other samples by the presence of a substitution of thymine for guanine at position 51, and adenine for thymine at position 203. The samples of the third population have guanine at position 121, while the other samples have either adenine or R (adenine/guanine), a similar situation is at position 402 (see Appendix A, Table A2).

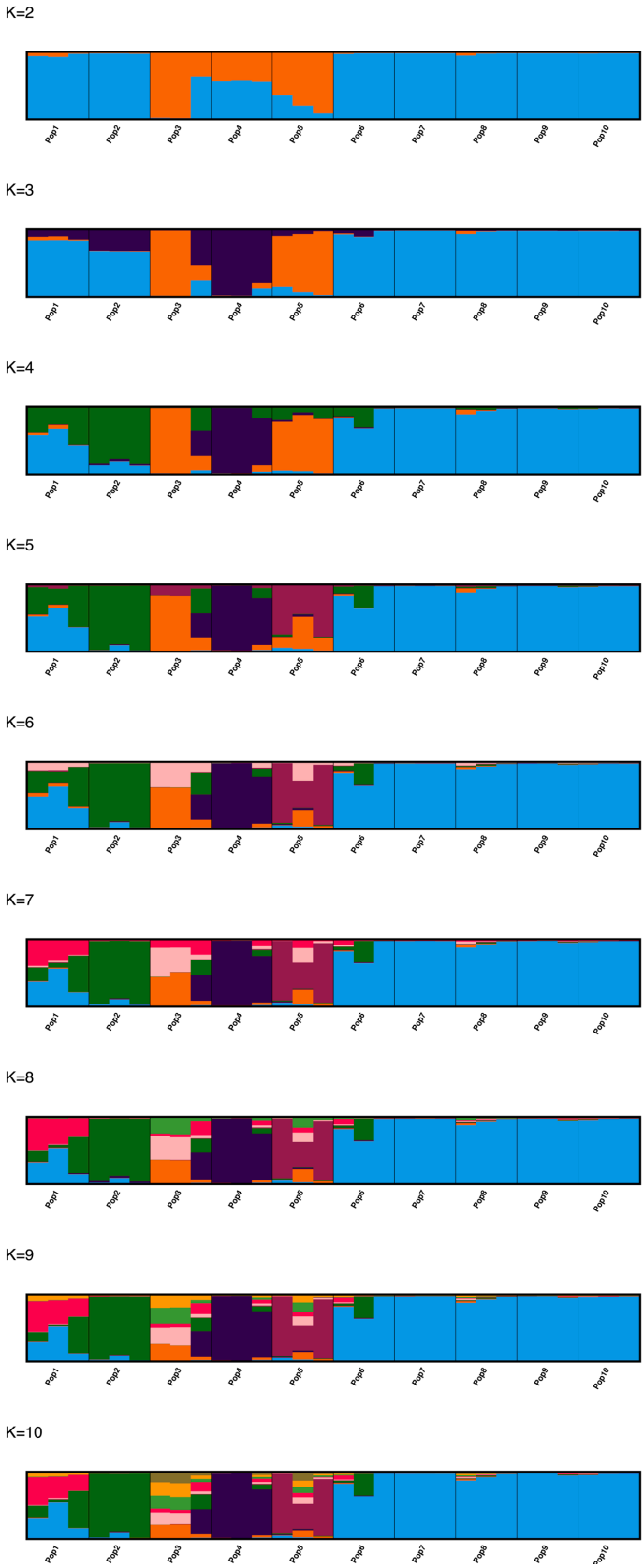

Figure 4. Genetic clustering of individuals from ten populations for different K values (from 2 to 10).

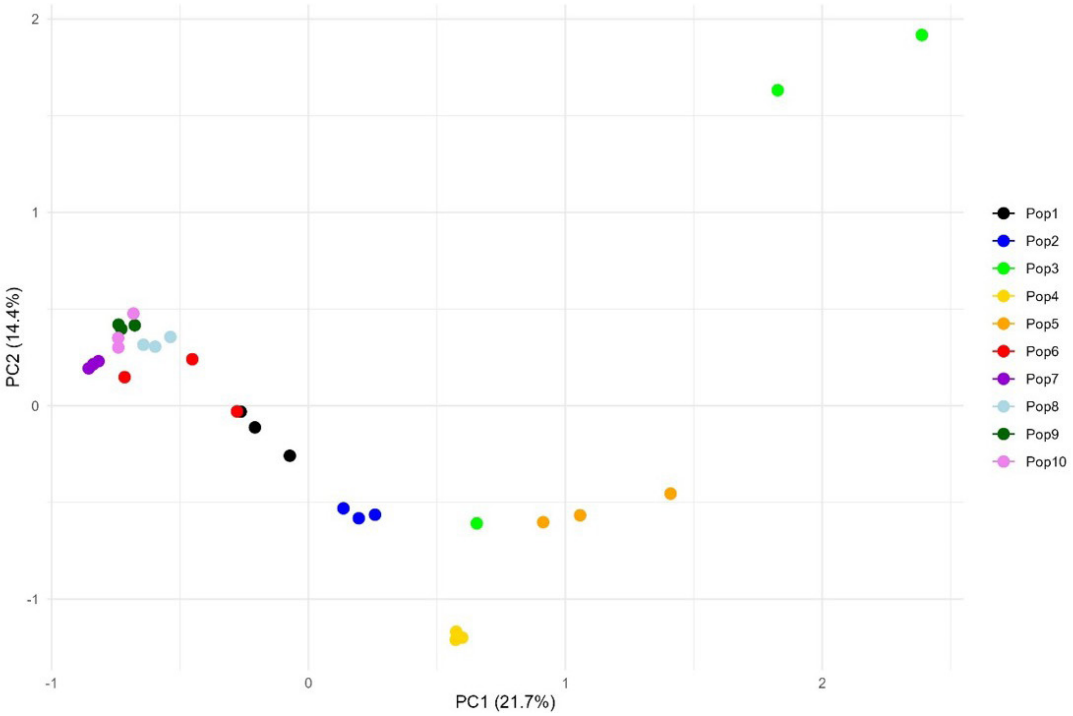

Figure 5. Genetic relationships among ten populations based on the first two principal components.

3.3. Molecular phylogeny

We sequenced nrITS by 13 plants of *A. sabulosum* from the gesamt distribution area and performed phylogenetic analysis with all ITS sequences of the subgenus *Allium* after Khassanov (2018) available in NCBI GenBank. There were 79 nrITS sequences from the subgenus *Allium* and sequences from *Allium ramosum* L., *A. tuberosum* L., *A. trifurcatum* and *A. oreoprasum* Schrenk. (*A.* subgen. *Butomissa* (Kamelin) N.Friesen), *A. mairei* H.Lév. and *A. cyathophorum* Bureau & Franch. (*A.* subgen. *Cyathophora* (R.M.Fritsch)R.M.Fritsch) and *A. tenuissimum* L., *A. austrosibiricum* N.Friesen (*A.* subgen. *Rhizirideum* (G.Don ex Koch) Wendelbo) has been chosen as the outgroup. The alignment length of 87 ITS sequences is 682 characters; 213 characters are constant, 87 are parsimony-informative, and 389 are parsimony-uninformative. Unweighted parsimony analysis resulted in two most parsimonious trees of 1414 steps (consistency index CI = 0.35; retention index RI = 0.78). AIC chose the substitution model HKY+G in jModelTest2 for the Bayesian analysis.

The resulting phylogenetic tree (Figure 6) divides the subgenus *Allium* into three monophyletic groups, where the first is closest to the outgroup, the clade with species of sections *Coerulea* (Omelczuk) F.O.Khass., *Haneltii* F.O.Khass. *A. macrostemon* (Sect. *Scorodon*?) and *Eremoprasa* (Kamelin) F.O.Khass., R.M.Fritsch & N.Friesen, with *Allium sabulosum* and *A. eremoprasum*. The second clade consists mainly of the monophyletic subclade with species of section *Codonoprasum* and two very small sections (sect. *Pallasii* and sect. *Kopetdagia*). The third clade unites all other sections of the subgenus *Allium* and is divided into two sister subclades: a smaller one with very weak support (PP 0.61) and consisting of *Allium margaritae* (sect. *Brewispatha*), the monotypic section *Mediasia* (*A. turkestanicum*), *A. setifolium* and *A. pamiricum* (sect. *Avulsea* after Khasanov 2018), and a larger, very well-supported clade with section *Allium* and many smaller sections (*Avulsa, Brevidentia, Crystalina*, and *Multicaulea*).

All sequences of *A. sabulosum* form strongly supported clades and are relatively clearly distributed among several subclades showing a geographical pattern, with some exceptions. The most distant are accession 4 from Kyzylorda region and accession 5 from Mangyshlak region.

In the rpl32-trnL spacer, the samples from the fourth and fifth populations, which are more distant to the other studied accessions, have two deletions. The first is 20 nucleotides in size (the other samples have the sequence ATAACACTTAGAAAACTTAG at this position). The second deletion in sample 5.1 is 11 nucleotides, and in sample 4.1, it is 6 nucleotides. In the resolution of the phylogenetic tree these deletions play no role (Figure 7). In the plastid tree, subgenus Allium is also divided into three clades, but in a slightly different constellation than in the nrITS tree. This could be the reason why we have far fewer sequences of the plastid fragment rpl32-trnL spacer in GenBank. Nevertheless, section *Eremoprasa* is in a clade with section *Coerulea*, section *Haneltia*, and *A. macrostemon*, just as in the nrITS tree.

3.4. Flow cytometry

The genome size (DNA content in nuclei) was determined from 10 populations of the studied species, the results of which are presented in Table 3.

The results obtained by flow cytometry showed polyploidy (diploid, tetraploid and hexaploid) of the species. Of all the populations, population 3 stood out, which shows different polyploidy (Table 3). In addition, hexoploidy is presented with a DNA content of 71.405 pg. Examples of histograms of the studied *Allium sabulosum* samples are shown in Figure 8.

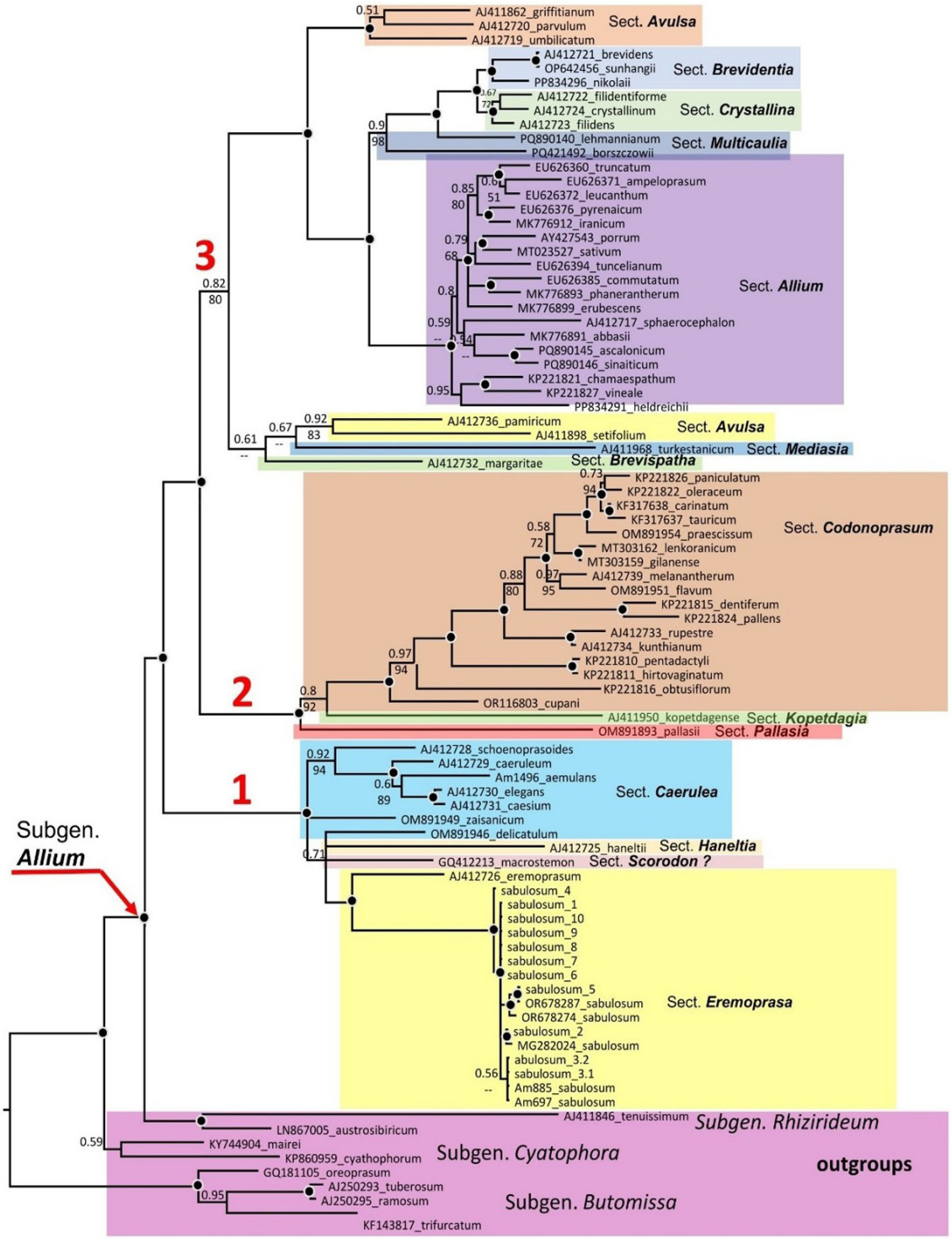

Figure 6. Phylogenetic nrITS tree of the subgenus *Allium* of genus *Allium*. Numbers by nodes represent Bayesian probabilities (above) and bootstrap support (1000 replicates). The joint presence of Bayesian probabilities over 0.98 and bootstrap support over 95% is indicated with a black dot. For the origin of samples without GenBank accession numbers, (see Appendix A, Table A3).

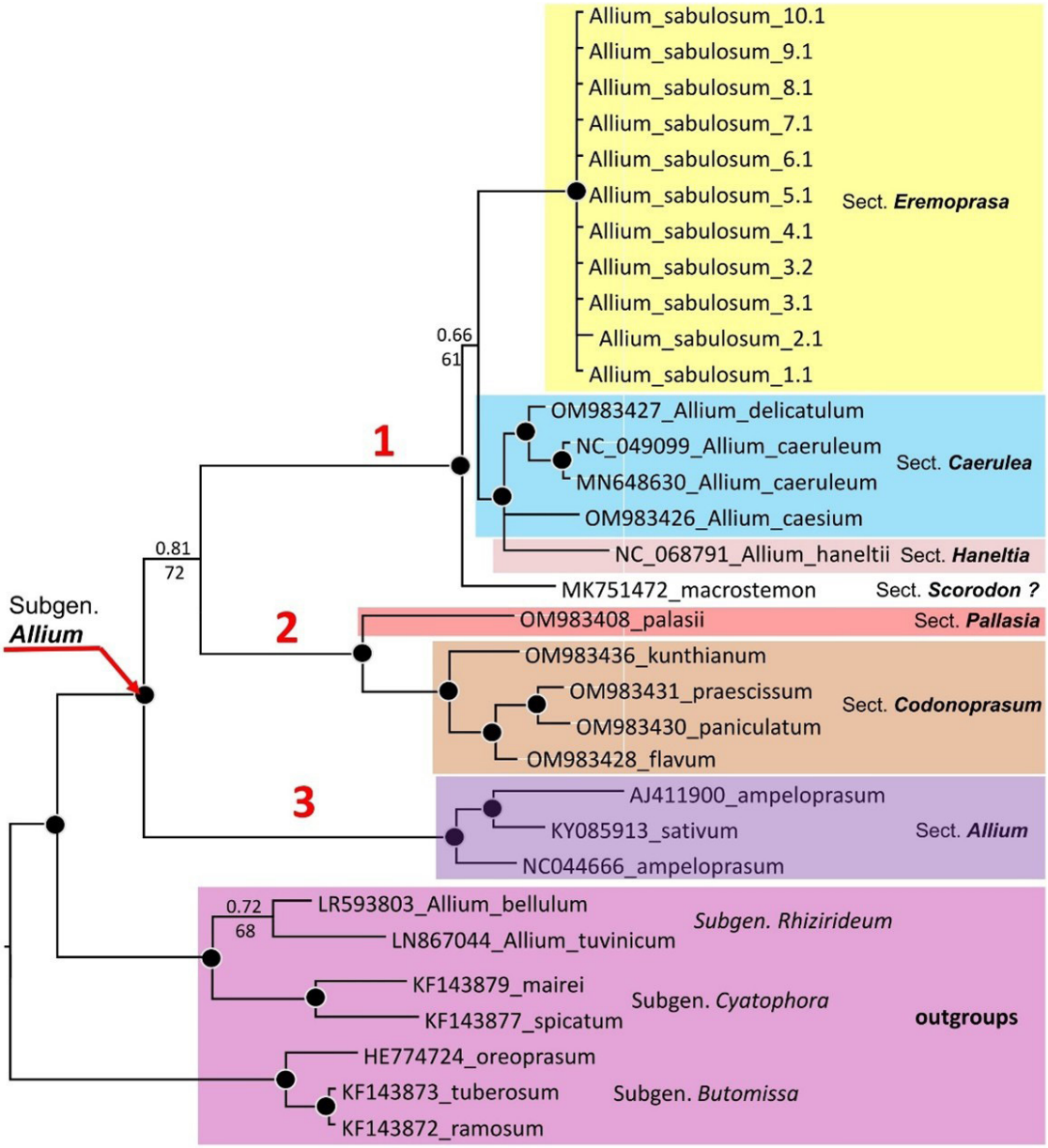

Figure 7. Phylogenetic plastide tree (rpl32-trnL) of the subgenus *Allium* of genus *Allium*. Numbers by nodes represent Bayesian probabilities (above) and bootstrap support (1000 replicates). The joint presence of Bayesian probabilities over 0.98 and bootstrap support over 95% is indicated with a black dot. For the origin of samples without GenBank accession numbers, (see Appendix A, Table A3).

Table 3. DNA content of the studied *Allium sabulosum* samples.

Cytotype	Population No.	Mean 2C ± SD, pg	CV	1Cx, Gbp	Index
2*x*	1 and 3	27,800 ± 0,299	1,08%	13,594	
4*x*	2 – 10	51,500 ± 1,119	2,17%	12,592	1,85
6*x*	3	71,405 ± 1,576	2,21%	11,639	2,57

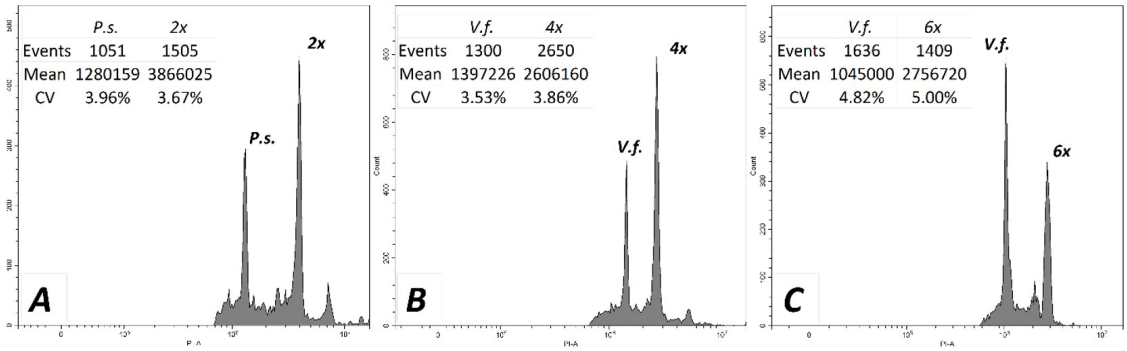

Figure 8. Examples of ungated flow cytometric histograms of the *Allium* samples (log scale). A – *Allium sabulosum* diploid cytotype; B – *Allium sabulosum* tetraploid cytotype; C – *Allium sabulosum* hexaploid cytotype. P.s. – *Pisum sativum* internal standard; V.f. – *Vicia faba* internal standard.

4. Discussion

The studied species has not been studied at the interpopulation level to date, there are only a few published ITS fragments (Abugalieva et al., 2017). According to our data, the studied species with the ITS fragment has a length of 643–645 bp, and the rpL32–trnL fragment is 838–873 bp. Also, the analysis of the *Allium sabulosum* ITS tree (Figure 6) showed a fairly clear geographic distribution of the studied species by collection sites. Samples taken from the Mangyshlak and Kyzylorda floristic region of the ITS tree differed from other populations, we associate this with the geographic distance.

According to the data (Vakhtina et al., 1977), the amount of DNA of the studied species is 21 pg, in our studies the average DNA content ranged from 27.8 pg to 71, 405 pg. In the cytometric study, population 3 stands out in particular, which shows different polyploidy (Table 3), but in the phylogenetic trees it was combined into one floristic region with other samples, which proves the presence of polyploidization of the species in the Balkhash-Alakol floristic region. Also, tetraploid samples from populations 6 to 9 in the Betpak-Dala floristic region were grouped together in two phylogenetic trees, we assume polyploidization of the species in this region as well.

According to the results of the research of many scientists on the genus (Xie et al., 2019, 2020; Friesen et al., 2021; Yusupov et al., 2021; Munavvarov et al., 2022), hybridization plays a role in the formation of *Allium* species. And in the dendrogram of population similarity based on iPBS-PCR analysis, especially against the background of all other studied samples, population 3 stands out, we assume that these samples are hybrids by distinguishing morphological features. In the phylogenetic position of the species in two types of trees (Figures 7 and 8), the supposed hybrid population is located next to other normal samples from the same floristic region. However, according to the authors (Munavvarov et al., 2022), the morphology of the species is not an important feature for study in the subgenus *Allium*, our studies show that in interpopulation analysis for *Allium sabulosum* it is important to take into account not only morphological features, but also the geographic distance when collecting samples. Unfortunately, due to the small number of *Allium sabulosum* populations collected in nature and analyzed, we cannot yet reliably explain hybridity and different polyploidy in populations. Additional karyological and population studies are needed here.

5. Conclusion

As a result of interpopulation study of psammophytic Iran–turanian species *Allium sabulosum* it was established that among the studied populations putative hybrid samples were found, which is confirmed by cytometric data. Only in the Balkhash-Alakol region, located on the eastern border of the range, have diploid, tetraploid, hexaploid forms, and in other regions only tetraploid forms were found. Finding diploid samples in the Balkhash-Alakol region may indicate the autochthonous origin of *Allium sabulosum* in this region. Unfortunately, the ITS sequence data and chloroplast fragments do not show this differentiation. For a reliable explanation of these changes, it is necessary to have a larger sample and further study of the species within the deserts of Kazakhstan (namely, the Balkhash-Alakol region).

Acknowledgements

This research is funded by the Science Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan Grant No. AP19679078 “Studying the species diversity of ecotone territory of north-eastern Betpakdala for preserving the relict gene pool of Kazakhstan arid flora” (2023-2025).

Data availability

All the data that support the findings of this study are available in the main text.

References

ABDILDANOV, D.S.H., VESSELOVA, P.V., KUDABAYEVA, G.M., OSMONALI, B.B., SKAPTSOV, M.V. and FRIESEN, N., 2025. Endemic of Kazakhstan *Allium lehmannianum* Merckl. ex Bunge and its Position within the genus *Allium*. *Plants*, vol. 14, no. 7, pp. 1113. http://doi.org/10.3390/plants14071113. PMid:40219181.

ABDULINA, S.A., 1999. *Spisok sosudistykh rasteniy Kazakhstana* [List of vascular plants in Kazakhstan]. Almaty, 52 p.

ABUGALIEVA, S., VOLKOVA, L., GENIEVSKAYA, Y., IVASCHENKO, A., KOTUKHOV, Y., SAKAUOVA, G. and TURUSPEKOV, Y., 2017. Taxonomic assessment of *Allium* species from Kazakhstan based on ITS and matK markers. *BMC Plant Biology*, vol. 17, suppl. 2, pp. 51-60. http://doi.org/10.1186/s12870-017-1194-0. PMid:29297332.

BAITENOV, M. S., 2001. *Flora of Kazakhstan: the genus complex of the flora*. vol. 2. 189 p.

CHASE, M.W., CHRISTENHUSZ, M.J.M., FAY, M.F., BYNG, J.W., JUDD, W.S. and SOLTIS, D.E., 2016. An update of the angiosperm phylogeny group classification for the orders and families of flowering plants: APG IV. *Botanical Journal of the Linnean Society*, vol. 181, no. 1, pp. 1-20. http://doi.org/10.1111/boj.12385.

CHERYOMUSHKINA, V.A., 2006. Populational strategy of *Allium* species (Alliaceae). In: D. IVANOVA, ed. *Plant, fungal and habitat diversity investigation and conservation: proceedings of IV Balkan Botanical Congress*. Sofia: Institute of Botany, Bulgarian Academy of Sciences, pp. 396-402.

CHOI, H.J. and OH, B.U., 2011. A partial revision of *Allium* (Amaryllidaceae) in Korea and north-eastern China. *Botanical Journal of the Linnean Society*, vol. 167, no. 2, pp. 153-211. http://doi.org/10.1111/j.1095-8339.2011.01166.x.

DARRIBA, D., TABOADA, G.L., DOALLO, R. and POSADA, D., 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods*, vol. 9, pp. 772. https://doi.org/10.1038/nmeth.2109.

DOLEŽEL, J., GREILHUBER, J., LUCRETTI, S., MEISTER, A., LYSÁK, M., NARDI, L. and OBERMAYER, R., 1998. Plant genome size estimation by flow cytometry: inter-laboratory comparison. *Annals of Botany*, vol. 82, suppl. 1, pp. 17-26. http://doi.org/10.1093/oxfordjournals.aob.a010312.

DOLEŽEL, J., SGORBATI, S. and LUCRETTI, S., 1992. Comparison of three DNA fluorochromes for flow cytometric estimation of nuclear DNA content in plants. *Physiologia Plantarum*, vol. 85, no. 4, pp. 625-631. http://doi.org/10.1111/j.1399-3054.1992.tb04764.x.

EARL, D.A. and VONHOLDT, B.M., 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, vol. 4, no. 2, pp. 359-361. http://doi.org/10.1007/s12686-011-9548-7.

EPIKTETOV, V., 2025 [viewed 10 June 2025]. *The genus Allium in the flora of Kazakhstan* [online]. Floristic list. Available from: https://www.plantarium.ru/ page/flora/id/1087.html. [In Russian]

EVANNO, G., REGNAUT, S. and GOUDET, J., 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, vol. 14, no. 8, pp. 2611-2620. http://doi.org/10.1111/j.1365-294X.2005.02553.x. PMid:15969739.

FELSENSTEIN, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution; International Journal of Organic Evolution*, vol. 39, no. 4, pp. 783-791. http://doi.org/10.1111/j.1558-5646.1985.tb00420.x. PMid:28561359.

FRIESEN, N., FRITSCH, R. and BLATTNER, F., 2006. Phylogeny and new intrageneric classification of *Allium* (Alliaceae) based on nuclear ribosomal DNA ITS sequences. *Aliso*, vol. 22, no. 1, pp. 372-395. http://doi.org/10.5642/aliso.20062201.31.

FRIESEN, N., HERDEN, T., LEWEKE, M., GRÜTZMACHER, L., FRAGMAN-SAPIR, O., HURKA, H. , BLATTNER, F.R. and FRITSCH, R.M., 2024. Dated phylogeny, phylogeography, and classification of *Allium* subgenus Amerallium (Amaryllidaceae) from the Old World, based on six DNA fragments. *Taxon*, vol. 73, no. 4, pp. 971-991. http://doi.org/10.1002/tax.13217.

FRIESEN, N., SMIRNOV, S.V., LEWEKE, M., SEREGIN, A.P. and FRITSCH, R.M., 2021. Taxonomy and phylogenetics of *Allium* section *Decipientia* (Amaryllidaceae): morphological characters do not reflect the evolutionary history revealed by molecular markers. *Botanical Journal of the Linnean Society*, vol. 197, no. 2, pp. 190-228. http://doi.org/10.1093/botlinnean/boab023.

FRIESEN, N., 2023. Introduction to Edible Alliums: Evolution, Classification and Domestication. In: H.D. RABINOWICH and B. THOMAS, eds. *Edible Allium: botany, production and uses*. Wallingford: CAB International, pp. 1-19.

FRITSCH, R.M. and FRIESEN, N., 2002. Evolution, domestication and taxonomy. In H.D. RABINOWITCH and L. CURRAH. *Allium* crop science: recent advances (pp. 5-30). Wallingford: CABI publishing. http://doi.org/10.1079/9780851995106.0005.

ILLUSTRATED IDENTIFIER OF PLANTS OF KAZAKHSTAN, 1969. *Nauka Kazahskoj SSR*. Alma-Ata: Kazakhstan, vol. 1. (In Russian).

ILLUSTRATED IDENTIFIER OF PLANTS OF KAZAKHSTAN, 1972. *Nauka Kazahskoj S.S.R.* Alma-Ata: Kazakhstan, vol. 2. (In Russian).

KALENDAR, R., ANTONIUS, K., SMYKAL, P. and SCHULMAN, A.H., 2010. iPBS: a universal method for DNA fingerprinting and retrotransposon isolation. *Theoretical and Applied Genetics*, vol. 121, no. 8, pp. 1419-1430. http://doi.org/10.1007/s00122-010-1398-2. PMid:20623102.

KAMELIN, R.V., 1973. Phylogenetic analysis of the natural flora of mountainous Central Asia Nauka, *Leningrad*, pp. 1-356. [In Russian].

KHASSANOV, F.O., 2018. Taxonomical and Etnobotanical aspect of *Allium* species from Middle Asia with reference to subgenus *Allium*. In M. SHIGYO, A. KHAR and M. ABDELRAHMAN, eds. *The Allium Genomes. Compendium of plant genomes.* Cham: Springer. http://doi.org/10.1007/978-3-319-95825-5_2.

KUMAR, S., STECHER, G., SULESKI, M., SANDERFORD, M., SHARMA, S., TAMURA, K., 2024. MEGA12: Molecular Evolutionary Genetic Analysis version 12 for adaptive and green computing. *Molecular Biology and Evolution*, vol. 41, no. 12, pp. msae263. http://doi.org/10.1093/molbev/msae263.

KUROCHKINA, L.Y.A., 1978. *Psammophytic vegetation of Kazakhstan*. Alma-Ata: Nauka, 272 p. [in Russian].

MAMYROVA, S., KUPRIYANOV, A., ISHMURATOVA, M., IVASHCHENKO, A., MYRZAGALIYEVA, A., ORAZOV, A. and KUBENTAYEV, S., 2025. The Current State of Populations of *Rhaponticum altaicum* (Asteraceae) in the Northern and Central Kazakhstan. *Diversity*, vol. 17, no. 3, pp. 206. http://doi.org/10.3390/d17030206.

METSALU, T. and VILO, J., 2015. ClustVis: a web tool for visualizing clustering of multivariate data using Principal Component Analysis and heatmap. *Nucleic Acids Research*, vol. 43, no. W1, pp. W566-W570. http://doi.org/10.1093/nar/gkv468. PMid:25969447.

MILLER, M.P., 1997. *Tools for population genetic analysis (TFPGA) 1.3: a windows program for the analysis of allozyme and molecular population genetic data*. USA: Department of Biological Sciences, Northern Arizona University.

MUNAVVAROV, A.; YUSUPOV, Z.; ERGASHOV, I.; TOJIBAEV, K. Sh.; DENG, T. and SUN, H., 2022. Complete chloroplast genomes of ten species from subgenus *Allium* (Allium, Amaryllidaceae). *Plant Diversity of Central Asia*, vol. 1, no. 2, pp. 67-81.

PAVLOV, N.V. and POLJAKOV, P.P., 1958. *Flora of Kazakhstan*. Alma-Ata: Academy of Sciences of the Kazakh SSR, Vol. 2. [In Russian].

PEAKALL, R. and SMOUSE, P.E., 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, vol. 6, no. 1, pp. 288-295. http://doi.org/10.1111/j.1471-8286.2005.01155.x.

PORRAS-HURTADO, L., RUIZ, Y., SANTOS, C., PHILLIPS, C., CARRACEDO, Á. and LAREU, M.V., 2013. An overview of STRUCTURE: applications, parameter settings, and supporting software. *Frontiers in Genetics*, vol. 4, pp. 98.

PRITCHARD, J.K., STEPHENS, M. and DONNELLY, M.S., 2000. Inference of population structure using multilocus genotype data. *Genetics*, vol. 155, no. 2, pp. 945-959. http://doi.org/10.1093/genetics/155.2.945. PMid:10835412.

RACHKOVSKAYA, E.I., VOLKOVA, E.A. and KHRAMTSOV, V.N., eds., 2003. *Botanical geography of kazakhstan and middle Asia (Desert Region)*. St. Petersburg, Russia: Komarov Botanical Institute of the Russian Academy of Sciences, pp. 424. (In Russian)

RONQUIST, F. and HUELSENBECK, J.P., 2003. MrBayes 3: bayesian phylogenetic inference under mixed models. *Bioinformatics*, vol. 19, no. 12, pp. 1572-1574. http://doi.org/10.1093/bioinformatics/btg180. PMid:12912839.

SAGALAYEV, V. A. and FIRSOV, G. A., 2014. Christian Steven (1781-1863) at the Lower Volga. *Memoranda Soc. Fauna Flora Fennica*, vol. 90, pp. 25-32.

SHAW, J., LICKEY, E.B., SCHILLING, E.E. and SMALL, R.L., 2007. Comparison of whole chloroplast genome sequences to choose non-coding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. *American Journal of Botany*, vol. 94, no. 3, pp. 275-288. https://doi.og/10.3732/ajb.94.3.275.

SILBIGER, V.N., HIRATA, M.H., LUCHESSI, A.D., GENVIGIR, F.D., CERDA, A., RODRIGUES, A.C., WILLRICH, M.A., ARAZI, S.S., DOREA, E.L., BERNIK, M.M., FALUDI, A.A., BERTOLAMI, M.C., SANTOS, C., CARRACEDO, A., SALAS, A., FREIRE, A., LAREU, M.V., PHILLIPS, C., PORRAS-HURTADO, L., FONDEVILA, M. and HIRATA, R.D., 2012. Differentiation of african components of ancestry to stratify groups in a case–Control study of a brazilian urban population. *Genetic Testing and Molecular Biomarkers*, vol. 16, no. 6, pp. 524-530. http://doi.org/10.1089/gtmb.2011.0267. PMid:22288895.

SKAPTSOV, M.V., KUTSEV, M.G., SMIRNOV, S.V., VAGANOV, A.V., UVAROVA, O.V. and SHMAKOV, A.I., 2024. Standards in plant flow cytometry: an overview, polymorphism and linearity issues. *Turczaninowia*, vol. 27, no. 2, pp. 86-104. http://doi.org/10.14258/turczaninowia.27.2.10.

SKVORTSOV, A.K., 1977. *Herbarium. Manual on methodology and technique*. Moscow: Nauka, 199 p. [In Russian]

SWOFFORD, D.L., 2002. *PAUP*: Phylogenetic Analysis Using Parsimony (*and Other Methods), version 4.0b10*. Sunderland, MA, USA: Sinauer Associates.

INTERNATIONAL PLANT NAMES INDEX – IPNI. THE ROYAL BOTANIC GARDENS KEW, THE HARVARD UNIVERSITY HERBARIA. THE AUSTRALIAN NATIONAL HERBARIUM and THE ROYAL BOTANIC GARDENS, 2025 [viewed 10 June 2025]. *Welcome to the International Plant Names Index (IPNI)* [online]. Available from: http://www.ipni.org

THE ROYAL BOTANIC GARDENS KEW, 2025 [viewed 10 December 2024]. *Plants of the World Online* [online]. Royal Botanic Gardens, Kew. Available from: http://www.plantsoftheworldonline.org

VAKHTINA, L.I., ZAKIROVA, R.O., VAKHTIN, Y.B., 1977. Interspecific differences in DNA content and taxonomically significant characters in the genus *Allium* L. (Liliaceae). *Botanicheskii Zhurnal*, vol. 262, pp. 677-684.

VVEDENSKY, A.I., 1971. *Genus Allium L. Identifier of plants of Central Asia*. Tashkent: FAN Uzbek SSR, vol. 2, pp. 39-89 [in Russian].

WHITE, T.J., BRUNS, T., LEE, S. and TAYLOR, J., 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: M.A. INNIS, D.H. GELFAND, J.J. SNINSKY and T.J. WHITE, eds. *PCR protoclos: a guide to methods and applications*. San Diego: Academic Press, pp. 315-322. http://doi.org/10.1016/B978-0-12-372180-8.50042-1.

XIE, D.F., TAB, J.B., YU, Y., GUI, L.J., SU, D.M., ZHOU, S.D. and HE, X.J., 2020. Insights into phylogeny, age, and evolution of *Allium* (Amaryllidaceae) based on the whole plastome sequences. *Annali di Botanica*, vol. 125, no. 7, pp. 1039-1055. http://doi.org/10.1093/aob/mcaa024.

XIE, D.F., YU, H.X., XIE, C., DENG, Y.Q., CHEN, Y.P., YU, Y., ZHOU, S.D. and HE, X.J., 2019. Phylogeny of *Allium* sect. *Daghestanica* and adaptive evolution of *Allium* (Allioideae) species revealed by the chloroplast complete genome. *Front. Plant Sci*, vol. 10, pp. 460. http://doi.org/10.3389/fpls.2019.00460. PMid:31114591.

YUSUPOV, Z., DENG, T., VOLIS, S., KHASSANOV, F., MAKHMUDJANOV, D., TOJIBAEV, K. and SUN, H., 2021. Phylogenomics of *Allium* section *Cepa* (Amaryllidaceae) provides new insights on domestication of onion. *Plant Diversity*, vol. 43, no. 2, pp. 102-110. http://doi.org/10.1016/j.pld.2020.07.008. PMid:33997542.

YUSUPOV, Z., ERGASHOV, I., VOLIS, S., MAKHMUDJANOV, D., DEKHKONOV, D., KHASSANOV, F. and SUN, H., 2022. Seed macro-and micromorphology in *Allium* (Amaryllidaceae) and its phylogenetic significance. *Annals of Botany*, vol. 129, no. 7, pp. 869-911. http://doi.org/10.1093/aob/mcac067. PMid:35696666.

Supplementary Material

Supplementary material accompanies this paper.
Appendix A. Molecular genetic analysis.
Table A1. Matrix of differentiation of Allium sabulosum populations
Table A2. Nucleotide substitutions in the ITS fragment of nuclear DNA of the studied A. sabulosum samples
Table A3. Voucher specimens of the studied species.
Appendix B. Structure analysis.
Figure B1. Logarithm of the probability of data in A. sabulosum populations
Figure B2. Genotyped A. sabulosum lines using the Evanno method
This material is available as part of the online article from https://doi.org/10.1590/1519-6984.298182