UDC 574.91 **Research article**

Spatial organization, DNA barcoding, and DNA sample bank establishment of wild populations of *Malus sieversii* for genetic resource conservation of an endangered species

Symbat S. Islamova[1,2], Almagul A. Kakimzhanova[2], Damira A. Dyussembekova[2] Zhanar S. Nurgozhanova[3], Bayan D. Sadakova[4], Aidana S. Nurtaza[2]

[1]NJSC "S. Seifullin Kazakh Agrotechnical Research University", Astana, Kazakhstan
[2]LLP "National Center for Biotechnology", Astana, Kazakhstan
[3]State National Nature Park "Zhongar-Alatau", Sarkand, Kazakhstan
[4]State National Nature Park "Tarbagatai", Urzhar village, Kazakhstan

Corresponding author: Aidana S. Nurtaza: nurtazaaidana6@gmail.com
Co-authors: (1: SI) islamovasimbat@gmail.com; (2: AK) kakimzhanova@biocenter.kz, (3: DD) dussembekova.damira@gmail.com, (4: ZhN) zhanar.sabitovna@bk.ru; (5: BS) tarbaga-tayy@mail.ru

Received: 10.07.2026 **Accepted:** 15.09.2026 **Published:** 30.09.2026

Copyright: Islamova S.S. et al. This article is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0/)

Abstract

Background and Aim. *Malus sieversii* (Ledeb.) *M. Roem.* is the primary wild progenitor of cultivated apple and is listed in the Red Book of the Republic of Kazakhstan as well as classified as Vulnerable by the IUCN. Conservation of its natural populations requires systematic monitoring and detailed characterization of their distribution and genetic resources. This study aimed to assess the spatial distribution, elevational patterns, and molecular identity of *M. sieversii* populations in the Zhongar-Alatau and Tarbagatai State National Nature Parks.

Materials and Methods. A total of 156 trees were surveyed. Geographic coordinates and elevation were recorded for each tree using GPS. The spatial structure of sampling points was analyzed using pairwise geographic distance matrices and visualized in QGIS. Elevational differences among sampling sites were evaluated using the Kruskal-Wallis test. Genomic DNA was extracted from leaf samples in triplicate to establish a DNA sample bank. Species identification was confirmed by DNA barcoding using the chloroplast markers trnH-psbA, matK, and rbcL.

Results. Geographic distance analysis revealed a hierarchical spatial structure of the sampled trees, comprising three spatial scales – local, subregional, and regional – that separated the Tarbagatai trees from those of Zhongar-Alatau. Elevational distribution differed significantly among sampling sites (Kruskal–Wallis test: $H = 74.10$, $p<0.001$). Trees in Zhongar-Alatau occurred at 1091–1461 m (1308 ± 104 m), whereas Tarbagatai trees occupied lower elevations (834–1119 m; 888 ± 87 m). DNA was successfully extracted from all 156 trees, and DNA barcoding of 11 representative samples confirmed species identity with 98–100% sequence similarity. All sequences were deposited in GenBank.

Conclusion. The geographic and ecological distinctiveness of the Tarbagatai population, together with the established DNA sample bank, provides a scientific basis for long-term monitoring and conservation of *M. sieversii* genetic resources in protected areas of Kazakhstan.

Keywords: *Malus sieversii*; wild populations; spatial distribution; elevational analysis; biodiversity monitoring; protected areas.

Introduction

Malus sieversii (Ledeb.) M. Roem. is a wild species of the genus *Malus*, which is the main ancestor of modern cultivated apple trees (M. × domestica Borkh.). The Central Asian origin of this species was first scientifically substantiated by *N.I. Vavilov* during expeditions in the 1920s and 1930s, when he

identified the mountainous regions of Kazakhstan as one of the world's centers of origin of cultivated plants [1, 2]. Academician *A.D. Dzhangaliyev* made an outstanding contribution to the study of this species, devoting more than 50 years to the comprehensive ecological and biological study of apple trees in Kazakhstan and the development of measures for their protection [3].

At the same time, over the past half century, the area of apple trees has decreased catastrophically: according to various estimates, from 70 to 80% of natural plantings in the vicinity of Almaty have been lost due to economic development of the territories [3]. The main threats to the species remain habitat fragmentation and degradation as a result of grazing, logging and urbanization, introgression of genes from cultivated varieties during cross-pollination, as well as damage to plantings by pests and diseases [4, 5].

Due to the unfavorable state of the population, *M. sieversii* is listed in the Red Book of the Republic of Kazakhstan and is also recognized as a vulnerable species in the IUCN Red List [6]. The species is included in the list of protected plants within state national natural parks in accordance with the Law of the Republic of Kazakhstan “On Specially Protected Natural Areas” (2006). Despite the existing protection system, many genetic reserves are experiencing degradation of stands, aging populations, and a decline in the proportion of young individuals, which threatens the natural regeneration of the species [7].

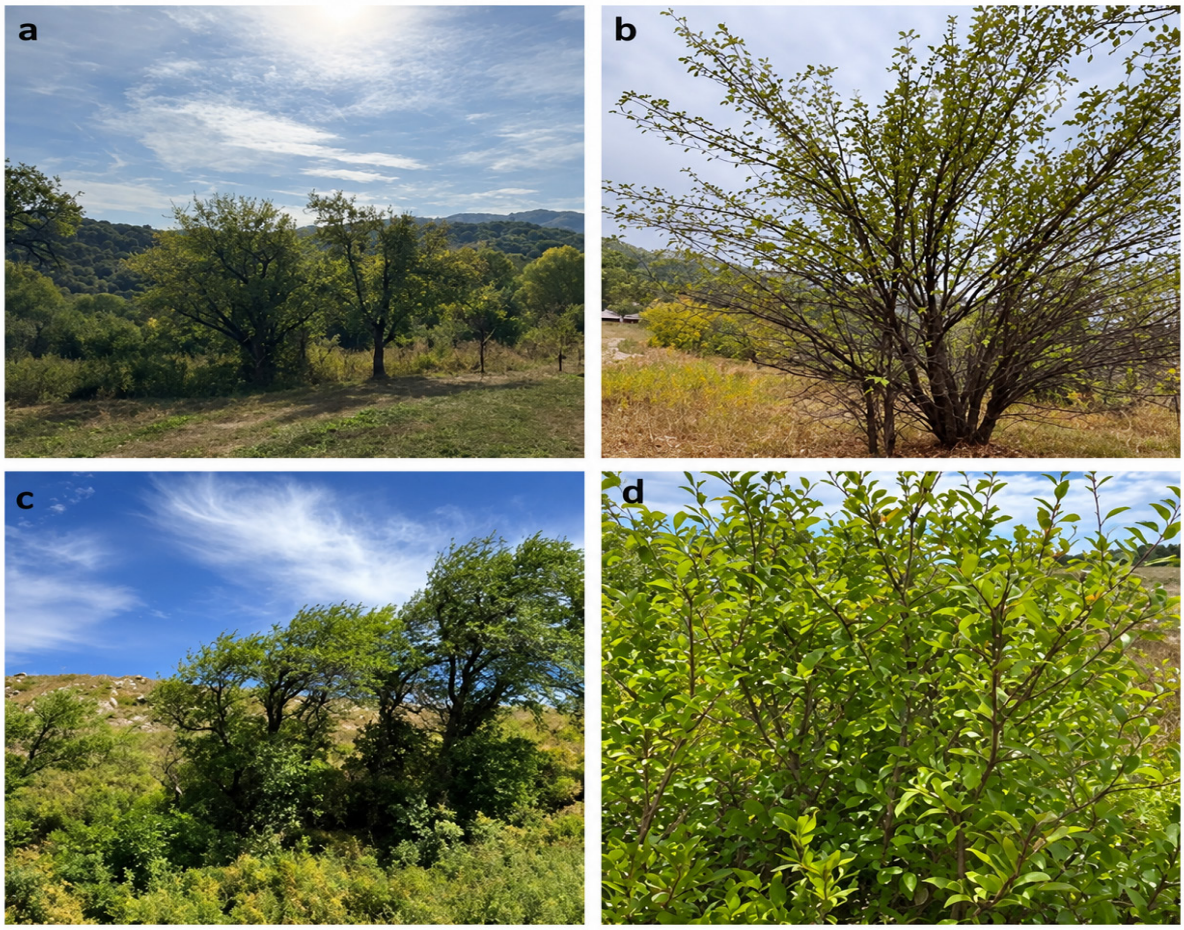
Systematic monitoring of natural populations is essential for developing scientifically sound conservation measures for *M. sieversii.* Spatial analysis of habitats allows us to identify patterns in the species' distribution, assess the degree of population fragmentation, and identify territorial units requiring priority attention [8]. Such data form the basis for planning environmental protection measures, organizing monitoring and rational management of genetic resources of wild fruit plants.

The aim of this study is a comprehensive characterization of wild populations of *M. sieversii* in the territory of the Zhongar-Alatau State National Park and the Tarbagatai State National Park, including an analysis of the spatial and elevational distribution based on expedition data and geoinformation analysis, as well as molecular genetic identification of samples using DNA barcoding and the formation of a DNA sample bank for the long-term preservation of the genetic resources of the endangered species.

Materials and Methods

Field research was conducted on the territory of the Zhongar-Alatau State National Nature Park (SNNP) and the Tarbagatai State National Nature Park (Figure 1). Within the Zhongar-Alatau State National Nature Park, sampling was carried out at three sites: the Alakol branch (Kokzharskoye forestry, trees 1–49), the Lepsinsky branch (Chernovskoye forestry, trees 50–91), and the Sarkansky branch (Topolevskoye forestry, trees 92–131). Additionally, trees 132–156 were collected on the territory of the Tarbagatai State National Nature Park. A total of 156 *M. sieversii* trees were examined.

Trees were selected at each site using a route-based method: routes were planned jointly with State National Nature Park staff, based on their knowledge of the locations of *M. sieversii* stands, and followed existing tracks and access paths within each forestry district; within each district, all accessible apple tree stands were surveyed on foot, and every tree encountered was recorded regardless of age or condition. No minimum distance threshold was imposed between sampled trees; within compact stands, neighbouring trees were recorded as little as 1–2 m apart. The surveyed stands within each site were spatially discrete, separated from one another by 0.5–16.7 km and reached via existing tracks and access routes; the survey therefore covered accessible apple tree stands identified along these routes rather than the complete area of the national parks. The combined length of the walking transects along which trees were recorded was approximately 3.0 km, distributed among 13 discrete stands (up to four within each of the three Zhongar-Alatau forestry districts and four within the Tarbagatai population), which together spanned approximately 213 ha (Kokzharskoye), 611 ha (Chernovskoye), 503 ha (Topolevskoye), and 2671 ha (Tarbagatai). Trees were identified as *M. sieversii* in the field with the assistance of State National Nature Park staff, who accompanied the survey team and indicated the locations of the species at each site. For each tree, the geographic coordinates (latitude, longitude) and elevation above sea level were recorded using a GPS receiver.



a, b – *M. sieversii* in the Zhongar-Alatau State National Park;
c, d – *M. sieversii* in the Tarbagatai State National Park

Figure 1. Collection of M. sieversii samples from wild populations in natural parks

The geographic coordinates of the sampling points were used for spatial visualization in the QGIS geographic information system (version 3.36), which made it possible to display the location of the points relative to the relief and boundaries of the study areas.

To quantify the spatial distribution of samples, a matrix of pairwise geographic distances between all sampling points was generated. Geodesic distances were calculated in Python using the GeoPy library. Heatmaps were constructed using the resulting matrix using the seaborn library: an aggregated matrix grouped by sites and a full matrix of distances between all points. To provide a generalized characterization of the spatial structure, average pairwise distances (M ± SD) between points in each pair of sites were calculated.

To characterize the elevational distribution of *M. sieversii*, the minimum, maximum, mean, and standard deviation (M ± SD) of elevation above sea level were calculated for each site. The statistical significance of differences in elevational ranges between sites was assessed using the nonparametric Kruskal-Wallis test. For pairwise comparisons, the Mann-Whitney test with Bonferroni correction for multiple comparisons was used (adjusted significance level $\alpha = 0.0083$). Data processing was performed using the Python libraries pandas, numpy, scipy, matplotlib, and seaborn.

DNA barcoding

DNA was extracted using the modified CTAB method (2% CTAB, 1.4 M NaCl, 0.2% 2-mercaptoethanol, 20 mM EDTA, 100 mM Tris–HCl, pH 8.0) [9]. The quality and quantity of extracted DNA were verified using an ultraviolet spectrophotometer (NanoDrop1000; Thermo Scientific, Wilmington, NC, USA), and each sample was diluted to 50 ng/µL in sterile distilled water. A PCR reaction mixture of 20 µL comprised the following: 3 µL of template DNA (50 ng), 2 µL PCR buffer (10 ×), 1 µL $MgCl_2$ (25 mM), 1.5 µL of both forward and reverse primers (10 pmol), 0.25 µL Dream Taq polymerase, 1 µL dNTPs (10 mM) and deionised distilled water. The primers were synthesised at the Laboratory for the Development of

Molecular Diagnostic Approaches at the National Center for Biotechnology. The PCR cycling conditions were as follows: 95 °C for 5 min for DNA denaturation; 35 cycles of 95 °C for 1 min, 52–58 °C for 1 min, and 72 °C for 1 min; and 72 °C for 10 min. The amplified PCR products were electrophoresed on a 1% agarose gel (1 × TBE buffer & 0.5 µg/mL ethidium bromide) to check for the presence or absence of bands.

Purification of the PCR products was performed using shrimp alkaline phosphatase (SAP) and ExoI (Thermo Scientific) enzymes. The reaction mixture (20 µL) contained 1 × SAP buffer, 10 µL of PCR product, 3 units of exonuclease ExoI, and 1 unit of SAP and was incubated at 37 °C for 30 min, followed by inactivation of the enzyme at 75 °C for 15 min. The initial incubation hydrolysed excess primers and dephosphorylated the nucleotides, and the second incubation inactivated the enzymes, minimising the loss of the PCR product so that further sequencing could be performed without further purification using the columns. DNA sequencing was performed using an ABI Prism BigDye Terminator reagent kit version 3.1 (Austin, TX, USA), followed by analysis of the reaction products on an automated sequencer (Applied Biosystems 3730 DNA Analyser, Tokyo, Japan). The nucleotide sequences of the analysed samples were retrieved and edited using SeqMan II (version 6.1; DNASTAR Inc., Madison, WI, USA).

Results

Field surveys showed that *M. sieversii* occurred throughout the investigated areas of Zhongar-Alatau and Tarbagatai State National Parks. The species was recorded both on slopes of varying steepness and in relatively flat areas, indicating that the surveyed trees were not restricted to a single topographic position. Nevertheless, their distribution within the study area was distinctly uneven.

Most of the recorded trees occurred either as isolated individuals or as small localized groups. Large continuous stands were virtually absent from the surveyed sites. The spatial distribution of the sampling points therefore had a fragmented and mosaic pattern. This pattern is also evident in Figure 2, where the sampling points form several spatially separated clusters corresponding to the investigated forestry districts.

Field observations further revealed differences in age structure among the four sites, assessed qualitatively in the field (young, mature, old); a quantitative demographic survey (e.g., diameter or height class distribution) was not conducted. Kokzharskoye contained both young and mature trees. Chernovskoye was dominated by mature and old trees. Topolevskoye, part of a protected gene reserve where livestock grazing is prohibited, contained the oldest trees recorded in the study, with apparently low natural regeneration. Tarbagatai, where livestock grazing is unregulated, contained mature trees alongside a comparatively large number of young individuals.

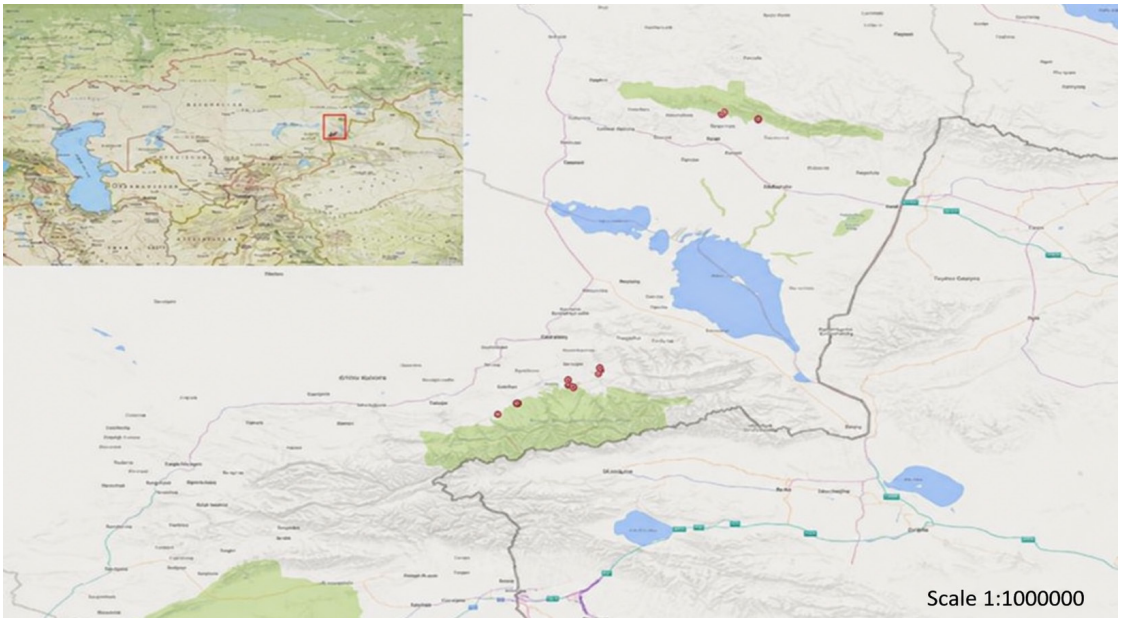


Figure 2. Spatial arrangement of sampling points on the territory of the Zhongar-Alatau State National Park and the Tarbagatai State National Park

The surveyed *M. sieversii* trees occupied a broad overall elevational range, extending from 834 to 1461 m above sea level. However, the elevation of the sampling points differed considerably between the two national parks.

In Zhongar-Alatau State National Park, the examined trees were recorded at elevations ranging from 1091 to 1461 m above sea level. The mean elevation was 1308 m, with a standard deviation of 104 m. Most sampling points in this national park were therefore located within the middle part of the total elevational range recorded during the study.

The elevational characteristics also differed among the three surveyed forestry districts of Zhongar-Alatau. In Kokzharskoye forestry, trees were recorded between 1266 and 1435 m above sea level, representing the narrowest elevational interval among the Zhongar-Alatau sites. In Chernovskoye forestry, the elevation ranged from 1209 to 1461 m, with the highest sampling point of the entire study recorded at this site.

Topolevskoye forestry occupied the broadest elevational interval among the Zhongar-Alatau sites. Trees were recorded from 1091 to 1457 m above sea level, giving a total elevational range of 366 m. This site also showed the greatest variation in elevation, with a standard deviation of 130 m. The minimum elevation recorded in Topolevskoye was 118 m lower than that in Chernovskoye and 175 m lower than that in Kokzharskoye. At the same time, the maximum elevations of all three Zhongar-Alatau sites were relatively similar and differed by no more than 26 m.

In Tarbagatai State National Park, the trees occurred at elevations of 834–1119 m above sea level. The mean elevation was 888 m, with a standard deviation of 87 m. Tarbagatai therefore occupied the lowest elevational position among all investigated sites. Its mean elevation was 420 m lower than the mean recorded for Zhongar-Alatau State National Park.

The maximum elevation recorded in Tarbagatai exceeded the minimum elevation of the Zhongar-Alatau dataset by only 28 m. Consequently, the elevational distributions of the two national parks showed only limited overlap at the boundary between the highest Tarbagatai and lowest Zhongar-Alatau sampling points. The difference was particularly pronounced when Tarbagatai was compared with Kokzharskoye and Chernovskoye, whose minimum elevations were 1266 and 1209 m, respectively.

Overall, the survey identified two main elevational patterns: a predominantly higher-elevation group represented by the three Zhongar-Alatau forestry districts and a lower-elevation group represented by Tarbagatai. A detailed summary of the elevational characteristics of each site is provided in Table 1.

Table 1. Elevation characteristics of *M. sieversii* habitats in the surveyed areas

Plot	Number of trees, pcs	Elevation range, m	M ± SD, m	Median, m
Kokzharskoye	49	1266–1435	1338 ± 65	1299
Chernovskoye	42	1209–1461	1331 ± 88	1329
Topolevskoye	40	1091–1457	1246 ± 130	1222
Tarbagatai	25	834–1119	888 ± 87	856
All trees	156	834–1461	1240 ± 185	–

Because the elevation data did not meet the assumptions required for parametric comparison, differences among the four sites were evaluated using the Kruskal–Wallis test. The analysis revealed highly significant differences in elevation among the surveyed sites ($H = 74.10$, $p < 0.001$). This result indicated that the sampling points from the four sites did not originate from identical elevational distributions.

Subsequent pairwise comparisons with Bonferroni correction showed that Kokzharskoye and Chernovskoye forestries did not differ significantly in elevation ($U = 956$, $p = 0.566$). This was the only nonsignificant pairwise comparison. Their similarity was also reflected in their overlapping elevational ranges of 1266–1435 m and 1209–1461 m, respectively.

All other pairwise comparisons revealed statistically significant differences ($p < 0.001$). Topolevskoye differed significantly from both Kokzharskoye and Chernovskoye, despite the overlap in their upper elevational limits. This distinction was associated with the broader elevational range of Topolevskoye and the occurrence of its sampling points at lower elevations.

Tarbagatai differed significantly from each of the three Zhongar-Alatau forestry districts ($p<0.001$). These comparisons confirmed that Tarbagatai represented a separate lower-elevation group within the surveyed dataset. Thus, both the descriptive statistics and nonparametric analysis demonstrated clear elevational differentiation among the study sites.

Spatial structure of sampling points

Pairwise geographic distances were calculated for all surveyed trees to quantify the spatial relationships within and among the sampling sites. The resulting distance matrix revealed a distinct block structure corresponding to Kokzharskoye, Chernovskoye, Topolevskoye, and Tarbagatai forestry districts (Figure 3).

The diagonal blocks of the matrix contained the smallest distance values and represented comparisons among trees sampled within the same site. Mean within-site distances ranged from 0.69 to 6.83 km. These values were substantially lower than the distances calculated between different sites, demonstrating that sampling points within each forestry district formed relatively compact spatial groups.

The small within-site distances were observed despite the fragmented occurrence of individual trees in the field. Thus, although trees were frequently represented by isolated individuals or small groups, the sampling points assigned to the same forestry district remained geographically closer to one another than to points from other districts.

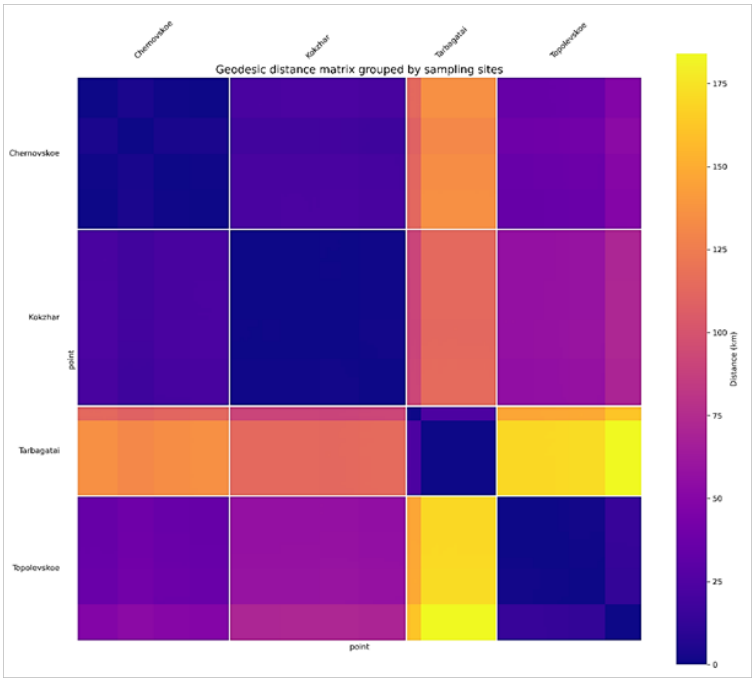


Figure 3. Matrix of pairwise geographic distances (km), grouped by sampling sites

The three Zhongar-Alatau forestry districts showed a gradual increase in pairwise geographic distances. The shortest inter-site distance was recorded between Kokzharskoye and Chernovskoye, with a mean value of approximately 20.87 km. Chernovskoye and Topolevskoye were separated by approximately 40 km, whereas the greatest distance within Zhongar-Alatau was recorded between Kokzharskoye and Topolevskoye, reaching approximately 60.90 km.

The distance between Kokzharskoye and Topolevskoye was therefore approximately three times greater than that between Kokzharskoye and Chernovskoye. Chernovskoye occupied an intermediate spatial position, being closer to Kokzharskoye than to Topolevskoye. This sequence was consistently reflected in both the aggregated and complete pairwise distance matrices.

The three forestry districts were distributed sequentially along the mountain system, with distances increasing progressively from Kokzharskoye through Chernovskoye to Topolevskoye.

Tarbagatai showed the greatest geographic separation from all other surveyed sites. Mean distances between Tarbagatai and the Zhongar-Alatau forestry districts ranged from 109.78 to 170.14

km. Even the shortest distance involving Tarbagatai was markedly greater than the maximum distance recorded among the three Zhongar-Alatau sites.

The maximum mean inter-site distance in the dataset was 170.14 km, recorded between Tarbagatai and the most distant Zhongar-Alatau site. These values clearly separated Tarbagatai from the cluster formed by Kokzharskoye, Chernovskoye, and Topolevskoye in the geographic distance matrix.

The contrast between within-site and between-region distances was especially pronounced. Within-site distances did not exceed a mean of 6.83 km, whereas distances between Tarbagatai and Zhongar-Alatau exceeded 100 km in all comparisons. Tarbagatai was therefore geographically isolated not only from individual forestry districts but from the entire group of Zhongar-Alatau sites.

Quantitative characteristics of the within- and between-site distances are presented in Table 2. The reported values represent mean distances calculated across all relevant pairs of sampling points, including the most distant points within and between the respective sites.

Analysis of the complete pairwise distance matrix was consistent with the patterns identified from the aggregated site-level data (Figure 4). Based on the magnitude of the pairwise distances, three spatial scales could be distinguished in the distribution of the surveyed trees; this grouping reflects the observed distance values rather than a formal statistical clustering analysis.

The first, local scale was represented by relatively compact clusters of sampling points within individual forestry districts. At this scale, the mean distances between trees were minimal and ranged from 0.69 to 6.83 km.

The second, subregional scale included the three forestry districts of Zhongar-Alatau State National Park. Their spatial arrangement was characterized by a consistent distance gradient of approximately 21–61 km, corresponding to their sequential positions along the mountain range.

The third, regional scale was represented by the separation of Tarbagatai from the Zhongar-Alatau sites. Mean distances at this level ranged from approximately 110 to 170 km and were considerably greater than distances observed within Zhongar-Alatau.

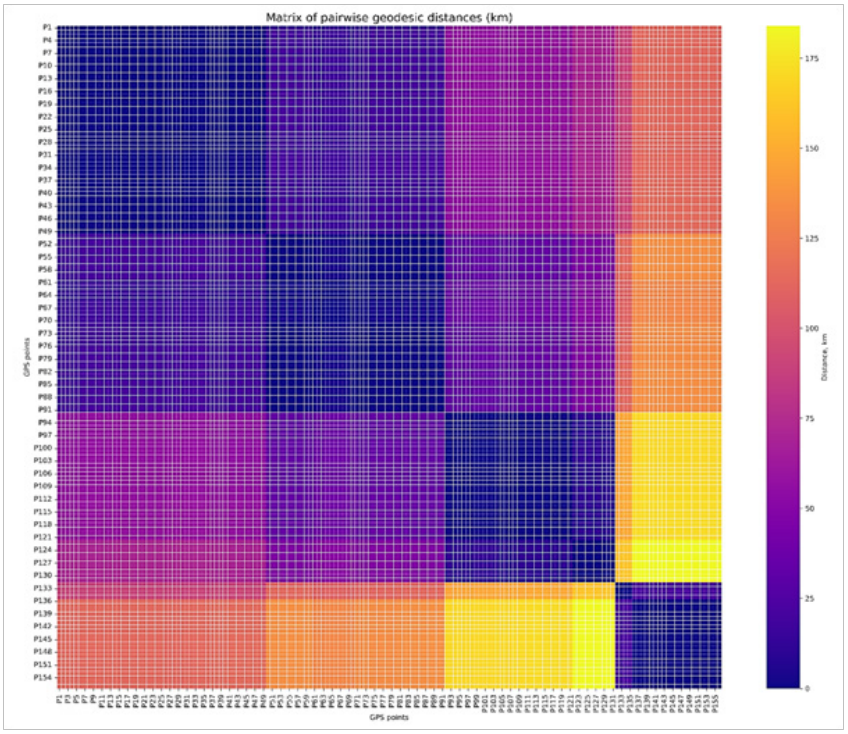


Figure 4. Complete matrix of pairwise geographic distances between all sampling points

Thus, the geographic distance analysis identified a structured rather than uniform distribution of the surveyed *M. sieversii* trees. The principal patterns included compact grouping within forestry districts, sequential spatial differentiation among the Zhongar-Alatau sites, and pronounced regional isolation of Tarbagatai.

Table 2. Average pairwise geographic distances (km) between sampling sites

Plot	Chernovskoye	Kokzharskoye	Tarbagatai	Topolevskoye
Chernovskoye	1.80	20.87	130.43	40.03
Kokzharskoye	20.87	0.69	109.78	60.90
Tarbagatai	130.43	109.78	6.83	170.14
Topolevskoye	40.03	60.90	170.14	5.61

DNA barcoding

To preserve the genetic resources of the wild *M. sieversii*, a DNA sample bank was created. DNA was extracted in triplicate from all 156 accessions collected during fieldwork using a modified CTAB method. The quality and concentration of the extracted DNA were monitored spectrophotometrically using a NanoDrop 1000 instrument (Thermo Scientific, USA); each sample was adjusted to a concentration of 50 ng/μL and stored at –20 °C in individually numbered tubes, each labelled with the tree identification number assigned during fieldwork, allowing every DNA sample to be traced back to its original tree and collection site. Thus, the DNA sample bank ensures long-term storage of genetic material from all surveyed *M. sieversii* populations from the Zhongar-Alatau and Tarbagatai State National Parks.

For molecular identification, DNA barcoding was performed using three chloroplast markers: rbcL, trnH-psbA, and matK, for 11 representative trees distributed across all four surveyed sites (three from Kokzharskoye, four from Chernovskoye, two from Topolevskoye, and two from Tarbagatai), together spanning most of the elevational range recorded at each site. These markers are standard for plant barcoding according to CBOL (Consortium for the Barcode of Life) recommendations and are widely used for species identification of wild fruit trees of the genus Malus.

The minimum and maximum length ranges (in base pairs, bp) for the trnH-psbA, matK, and rbcL regions were 231–420, 545–846, and 674–889, respectively. The average lengths of each region were 323.8, 758.2 and 766.5 bp, respectively. The GC content (%) for the trnH-psbA, matK, and rbcL regions ranged from 19.9 to 32.0%, 33.9 to 35.5%, and 42.3 to 43.7%, respectively. The mean GC values were 25.4%, 34.4%, and 42.6%, respectively (individual values for each of the 11 trees are given in Table 4). These GC content data are consistent with the known compositional properties of chloroplast markers in angiosperms: the trnH-psbA intergenic spacer is characteristically AT-rich, reflecting its non-coding nature, while the rbcL coding region shows substantially higher GC content, and matK occupies an intermediate position. All nucleotide sequences were manually edited using SeqMan II (version 6.1; DNASTAR Inc., Madison, WI, USA) to eliminate possible errors and improve the accuracy of the data. We used the BLAST method in the NCBI database to analyse the similarity of the obtained sequences; for each sequence, the top-ranking hit (best match) in the BLAST output was used to assess the degree of similarity with known reference sequences. The results showed a high level of similarity to the top hit, ranging from 98 to 100%, for all three markers. The sequencing results were uploaded to the international NCBI GenBank database, where each sample was assigned accession numbers for each of the three markers.

The resulting nucleotide sequences were uploaded to the international NCBI (National Center for Biotechnology Information) GenBank database, and accession numbers were assigned for the rbcL, matK and trnH markers (Table 3). This created a collection of nucleotide sequences, contributing to the global database of *M. sieversii* genetic resources in protected areas of Kazakhstan.

Table 3. NCBI GenBank Accession Numbers of rbcL, matK and trnH Sequences of *M. sieversii*

Sample	Accession Numbers		
	rbcL	matK	trnH
1	PZ462615	PZ462623	PZ469479
2	PZ462616	PZ462624	PZ469480
3	PZ452034	PZ462625	PZ469481
4	PZ452035	PZ462626	PZ469482
5	PZ452036	PZ462627	PZ469483
6	PZ452037	PZ462628	PZ469484

Continuation of Table 3

7	PZ452038	PZ462629	PZ469485
8	PZ452039	PZ462630	PZ469486
9	PZ452040	PZ462631	PZ469487
10	PZ452041	PZ462632	PZ469488
11	PZ452042	PZ462633	PZ469489

Table 4. Sequence length and GC content of the trnH-psbA, matK, and rbcL markers for the 11 genotyped *M. sieversii* trees, by sampling site (accession numbers given in Table 3)

Sample	Site	Elevation, m	trnH-psbA		matK		rbcL	
			bp	GC, %	bp	GC, %	bp	GC, %
1	Tarbagatai	1116	270	20.0	838	34.0	707	42.3
2	Tarbagatai	834	268	20.1	839	34.2	860	42.3
3	Kokzharskoye	1277	247	19.4	846	33.9	889	42.3
4	Kokzharskoye	1293	253	19.8	577	35.0	718	42.6
5	Kokzharskoye	1431	231	19.9	832	34.3	674	42.4
6	Chernovskoye	1212	291	22.0	838	34.0	863	42.6
7	Chernovskoye	1461	221	33.0	832	34.3	684	42.4
8	Chernovskoye	1457	253	23.7	539	35.1	726	42.4
9	Chernovskoye	1375	275	20.4	818	34.1	851	42.5
10	Topolevskoye	1231	294	21.8	813	34.1	712	42.4
11	Topolevskoye	1454	278	21.9	542	35.2	857	42.7

Discussion

Elevational distribution and ecological differentiation

The elevational distribution of M. sieversii recorded in Zhongar-Alatau State National Park is consistent with the established ecological range of the species. According to *A.D. Dzhangaliev*, wild apple forests in Kazakhstan occur mainly at elevations of 1000–1800 m, with the most favorable conditions found between 1200 and 1500 m [3]. Comparable elevations of 1000–1500 m have been reported by Yan et al. [10] for populations in *Xinjiang*, China, while *Richards* et al. [11] recorded M. sieversii at 1190–1760 m in the Dzungarian Range (Zhongar-Alatau) of Kazakhstan. The elevational ranges of Kokzharskoye and Chernovskoye forestries therefore fall within the recognized ecological optimum of the species.

The wider elevational range and greater variability observed in Topolevskoye forestry may reflect the pronounced topographic and environmental heterogeneity of the Sarkansky branch of the Zhongar-Alatau. *Zhang* et al. [12] identified temperature regime, particularly annual temperature range and seasonal variability, as one of the principal environmental factors determining the distribution of *M. sieversii* in the mountains of Central Asia. The variation observed in Topolevskoye may therefore be associated with differences in temperature and other environmental conditions along the local elevation gradient, although this hypothesis requires direct analysis of climatic and habitat variables.

The Tarbagatai population occupied a markedly lower elevational zone than the Zhongar-Alatau populations. Its observed range closely corresponds to that reported by *Richards* et al. [11] for the same region (890–1120 m). Those authors characterized the area as a dry continental forest exposed to winter temperatures as low as −40 °C and identified the Tarbagatai population as a distinct genetic cluster. The statistically significant elevational separation detected in the present study thus supports the ecological distinctiveness of this population.

The low-elevation position of the Tarbagatai population may also increase its vulnerability to climate change. Predictive models developed by *Zhu* et al. [8] indicate that suitable habitats for *M. sieversii* are likely to decline particularly in areas below 1000 m. Because a substantial proportion of

the surveyed Tarbagatai trees occurred below this threshold, this population may face a comparatively high risk of habitat contraction. The identification of Tarbagatai as an inadequately protected region by *Zhu* et al. [8] further emphasizes the need for targeted monitoring and conservation measures.

Population condition and fragmented distribution

Age structure differed among sites. Topolevskoye - part of a protected gene reserve where livestock grazing is excluded - contained the oldest trees recorded in this study and showed the lowest apparent regeneration, consistent with the findings of *Bekova* et al. [7], who identified stand degradation and senescence as major threats to *M. sieversii* populations in Zhongar-Alatau State National Park. In contrast, Tarbagatai, where livestock grazing is unregulated, contained a comparatively large number of young trees alongside mature individuals, suggesting more active regeneration despite the absence of formal protection from grazing; Kokzharskoye and Chernovskoye showed intermediate patterns. This contrast between a strictly protected but poorly regenerating site (Topolevskoye) and an unprotected but apparently more actively regenerating population (Tarbagatai) runs counter to the assumption that formal protection alone ensures successful natural regeneration, and is consistent with *Wang* et al. [13], who highlighted population aging and habitat degradation as continuing threats to wild apple populations in Kazakhstan despite formal protection.

The fragmented occurrence of isolated trees and small groups observed in the present study is not restricted to the surveyed forestry districts. Based on 8,344 occurrence records, *Zhu* et al. [8] estimated that the current distribution of *M. sieversii* in Central Asia is highly fragmented and covers approximately 49,956 km^2. *Zhang* et al. [12] likewise described the species as discontinuously distributed in mountain valleys of the arid regions of Central Asia and attributed this structure partly to climatic changes since the late Miocene. For Xinjiang populations, *Yan* et al. [10] reported a similarly discontinuous distribution confined to four principal regions. The spatial pattern observed in the present study therefore appears to reflect a broader characteristic of the species shaped by both historical processes and contemporary environmental conditions.

Fragmentation may restrict pollen and seed dispersal between local groups, potentially reducing gene flow and increasing genetic differentiation. However, geographic fragmentation alone does not demonstrate genetic isolation. Molecular genetic analyses are required to determine whether the spatially separated groups identified in this study also represent genetically distinct populations.

Spatial organization and conservation implications

The pairwise distance analysis revealed a hierarchical spatial structure spanning local, subregional, and regional scales (see Results).

This pattern is broadly consistent with the geographic structuring reported for *M. sieversii* in previous genetic studies. *Zhang* et al. [12] described a gradual east-to-west replacement of genetic lineages along the Tien Shan. *Yan* et al. [10] also found a significant relationship between genetic and geographic distances among *M. sieversii* populations. In their study, 83.1% of genetic variation occurred within populations and 16.9% among populations, indicating that substantial diversity is retained locally despite regional differentiation.

The pronounced spatial separation of Tarbagatai is particularly noteworthy. *Richards* et al. [11], based on 949 trees from eight sites in Kazakhstan, demonstrated geographic structuring among wild apple populations and identified Tarbagatai as genetically differentiated. *Volk* et al. [14] similarly reported that genetic differentiation among *M. sieversii* collection sites was associated with geographic location, while southern and southwestern populations showed greater admixture. Together with these previous genetic findings, the spatial isolation documented here supports the treatment of Tarbagatai as a distinct priority for monitoring.

Nevertheless, the present geographic analysis alone is insufficient to designate the four surveyed sites as independent conservation units. Such designation would require supporting evidence from population-genetic, demographic, and ecological analyses. At this stage, the four sites can more appropriately be regarded as separate monitoring units. Differentiated monitoring would make it possible to evaluate regeneration, stand condition, habitat change, and potential loss of connectivity within each area.

Overall, the results demonstrate that the surveyed *M. sieversii* populations are characterized by fragmented occurrence, uneven regeneration, hierarchical spatial organization, and pronounced

regional separation of Tarbagatai. These findings support the need for long-term site-specific monitoring and targeted in situ conservation, particularly in low-elevation and geographically isolated populations. In practical terms, priority should be given to regular monitoring of natural regeneration and stand condition at all four sites, to maintaining or restoring landscape connectivity between the fragmented stands to support pollen and seed dispersal, and to safeguarding the Tarbagatai population as a distinct conservation and monitoring priority given its combined geographic, elevational, and climatic vulnerability. The established DNA sample bank provides a practical basis for these efforts, enabling future genetic verification of restoration material, ex situ propagation, and, together with the ongoing project on disease resistance in wild Malus species, the selection of resistant genotypes for conservation breeding.

Conclusion

The study allowed us to characterize the spatial distribution and elevational range of wild *M. sieversii* (Ledeb.) M. Roem. in the Zhongar-Alatau and Tarbagatai State National Parks. During the fieldwork, 156 trees were examined at four sites, recording GPS coordinates and elevation.

M. sieversii populations are characterized by a fragmented hierarchical distribution: from compact local groups within forestry units (0.69–6.83 km) to a subregional gradient between the Zhongar-Alatau forestry units (21–61 km) and the Tarbagatai regional isolation (110–170 km). Elevation analysis revealed statistically significant differences between the sites (Kruskal-Wallis test: $H = 74.10$; $p < 0.001$): the Zhongar-Alatau forestry areas occupy higher elevations (1091–1461 m, M ± SD = 1308 ± 104 m), while Tarbagatai occupies a significantly lower elevational range (834–1119 m, M ± SD = 888 ± 87 m), which emphasizes not only its geographical but also its ecological isolation.

DNA was extracted from all 156 accessions in triplicate, and a DNA sample bank was created. DNA barcoding using the trnH-psbA, matK, and rbcL markers confirmed the species identity of all 11 analyzed samples (98–100% similarity to NCBI GenBank); the resulting sequences were deposited in GenBank and assigned accession numbers.

Age structure varied among sites, from the oldest trees and lowest apparent regeneration in the strictly protected, grazing-free Topolevskoye gene reserve to a comparatively large proportion of young trees in freely grazed Tarbagatai, highlighting the need for site-specific regeneration monitoring and restoration measures, particularly at Topolevskoye. The combination of field data, the DNA sample bank, and the nucleotide sequence collection forms a comprehensive basis for long-term monitoring and conservation of *M. sieversii* genetic resources in protected areas of Kazakhstan.

Authors' Contributions

AN: conceived and designed the study and acquired funding. SI and AK: developed the methodology and contributed to the interpretation of results, supervised the research. ZhN and BS: collected and prepared field samples. DD and AA: performed laboratory experiments, including DNA extraction, PCR amplification, and sequencing. AN and SI: conducted spatial analysis and data visualization. All authors contributed to the discussion, writing, and revision of the manuscript and approved the final version.

Information of funding

This research is funded by the Ministry of Science and Higher Education of the Republic of the Republic of Kazakhstan for 2025–2027 No. AP26102807 «Study of wild and introduced populations of the endangered species *Malus niedzwetzkyana* and *Malus sieversii* for the selection of disease-resistant genotypes and the creation of in vitro collection for conservation».

References

1 Vavilov, N.I. (1926). *Centry proishozhdeniya kul'turnyh rastenii* [*Centers of origin of cultivated plants*]. L.: VIR, 248. [*In Russian*].

2 Cornille, A., Gladieux, P., Smulders, M.J.M., et al. (2012). New insight into the history of domesticated apple: secondary contribution of the European wild apple to the genome of cultivated varieties. *PLOS Genetics*, 8(5), e1002703.

3 Dzhangaliev, A.D. (2003). The Wild Apple Tree of Kazakhstan. *Horticultural Reviews*, 29, 63–303.

4 Forsline, P.L., Aldwinckle, H.S., Dickson, E.E., et al. (2003). Wild Apples and Fruit Trees of Central Asia. *Horticultural Reviews*, 29, 1–61.

5 Poyraz, I., Janisiewicz, W.J., Saftner, R., *et al.* (2021). Genetic Admixture in the Population of Wild Apple (Malus sieversii) from the Tien Shan Mountains, Kazakhstan. *Diversity*, 13(1), 29.

6 IUCN. (2007). Malus sieversii. The IUCN Red List of Threatened Species 2007: e.T32363A9693009. https://doi.org/10.2305/IUCN.UK.2007.RLTS.T32363A9693009.en

7 Bekova, R., Wennmann, J.T., Peil, A., *et al.* (2025). Assessment of genetic diversity and population structure of Malus sieversii and Malus niedzwetzkyana from Kazakhstan using high-throughput genotyping. *Tree Genetics & Genomes*, 21.

8 Zhu, Y., Luo, X., Watcharamongkol, S., et al. (2022). Wild Apples Are Not That Wild: Conservation Status and Potential Threats of Malus sieversii in the Mountains of Central Asia Biodiversity Hotspot. *Diversity*, 14(6), 489.

9 Doyle, J. (1991). *DNA protocols for plants. In Molecular Techniques in Taxonomy*. (eds Hewitt, G. M. et al.) Springer, NATO ASI Series, 57 283–293.

10 Yan, G., Long, H., Song, W., Chen, R. (2008). Genetic polymorphism of *Malus sieversii* populations in Xinjiang, China. *Genetic Resources and Crop Evolution*, 55, 171–181.

11 Richards, C.M., Volk, G.M., Reilley, A.A., et al. (2009). Genetic diversity and population structure in *Malus sieversii*, a wild progenitor species of domesticated apple. *Tree Genetics & Genomes*, 5, 339–347.

12 Zhang, H.X., Wen, Z.B., Wang, Q. (2022). Population genetic structure of *Malus sieversii* and environmental adaptations. *Chinese Journal of Plant Ecology*, 46(9), 1098–1108.

13 Wang, N., Jiang, S., Zhang, Z., et al. (2018). Malus sieversii: the origin, flavonoid synthesis mechanism, and breeding of red-skinned and red-fleshed apples. *Horticulture Research*, 5,70.

14 Volk, G.M., Henk, A.D., Richards, C.M., Forsline, P.L., Chao, C.T. (2013). *Malus sieversii:* A Diverse Central Asian Apple Species in the USDA-ARS National Plant Germplasm System. *HortScience*, 48(12), 1440–1444.

Жойылып кету қаупі төнген *Malus sieversii* түрінің генетикалық ресурстарын сақтау мақсатында табиғи популяцияларының кеңістіктік ұйымдасуын зерттеу, ДНҚ-баркодтау және ДНҚ банкін қалыптастыру

Исламова С.С.[1,2], Какимжанова А.А.[2], Дюсембекова Д.А.[2], Нұрғожанова Ж.С.[3], Садакова Б.Д.[4] Нұртаза А.С.[2]

[1]«С.Сейфуллин атындағы Қазақ агротехникалық зерттеу университеті» КеАҚ
Астана, Қазақстан
[2]«Ұлттық биотехнология орталығы» ЖШС, Астана, Қазақстан
[3]«Жоңғар-Алатауы» мемлекеттік ұлттық табиғи паркі, Сарқан, Қазақстан
[4]«Тарбағатай» мемлекеттік ұлттық табиғи паркі, Үржар ауылы, Қазақстан

Түйін

Алғышарттар мен мақсат. *Malus sieversii* (Ledeb.) M. Roem. – мәдени алма сорттарының негізгі арғы тегі болып табылатын жабайы өсетін түр. Ол Қазақстан Республикасының Қызыл кітабына енгізілген және IUCN жіктемесі бойынша осал түр ретінде танылған. Бұл түрдің табиғи популяцияларын сақтау жүйелі мониторинг жүргізуді және олардың өсу орындарын жан-жақты сипаттауды талап етеді. Осы жұмыста «Жоңғар-Алатауы» және «Тарбағатай» мемлекеттік ұлттық табиғи парктері аумағында жүргізілген экспедициялық зерттеулердің нәтижелері ұсынылған. Зерттеудің мақсаты – *M. sieversii* популяцияларының кеңістіктік таралуын, өсуінің биіктік жағдайларын және молекулалық-генетикалық сипаттамаларын анықтау.

Материалдар мен әдістер. Зерттеу барысында 156 ағашқа зерттеу жүргізілді. Әрбір үлгі үшін GPS көмегімен географиялық координаттары мен теңіз деңгейінен биіктігі тіркелді.

Популяциялардың кеңістіктік құрылымы жұптық географиялық қашықтықтар матрицасы негізінде бағаланып, кейін QGIS бағдарламасында геоақпараттық өңдеу және визуализация жүргізілді. Зерттеу учаскелері арасындағы биіктік диапазондарының айырмашылығы Краскел-Уоллис параметрлік емес критерийі арқылы талданды. Жапырақ үлгілерінен үш қайталымда геномдық ДНҚ бөлініп алынды, оның негізінде ДНҚ үлгілерінің банкі қалыптастырылды. Түрлік сәйкестендіру хлоропласттық trnH-psbA, matK және rbcL маркерлері бойынша ДНҚ-баркодтау әдісімен расталды.

Нәтижелер. Кеңістіктік талдау зерттелген ағаштардың үш деңгейден тұратын иерархиялық кеңістіктік құрылымын көрсетті: жергілікті (орманшылықтар шегінде), субөңірлік (Жоңғар-Алатауы орманшылықтары арасында) және өңірлік (Тарбағатай популяциясының оқшауланууы). Биіктік талдауы зерттеу учаскелері арасындағы өсу биіктігінің статистикалық тұрғыдан мәнді айырмашылықтарын анықтады (Краскел–Уоллис тесті: H = 74,10; p<0,001): Жоңғар-Алатауы популяциялары неғұрлым жоғары биіктік белдеуінде (1091–1461 м, M ± SD = 1308 ± 104 м), ал Тарбағатай популяциясы неғұрлым төменгі биіктік белдеуінде (834–1119 м, M ± SD = 888 ± 87 м) таралған. Барлық 156 үлгіден үш қайталымда ДНҚ бөлініп алынып, ДНҚ үлгілерінің банкі қалыптастырылды. trnH-psbA, matK және rbcL маркерлері бойынша жүргізілген ДНҚ-баркодтау талданған 11 үлгінің түрлік сәйкестігін (98–100%) растады, ал алынған нуклеотидтік тізбектер NCBI GenBank дерекқорына тіркеу нөмірлері беріліп енгізілді.

Қорытынды. Алынған нәтижелер Тарбағатай популяциясының географиялық және экологиялық оқшауланғанын дәлелдейді. Қалыптастырылған ДНҚ банкімен бірге бұл деректер Қазақстанның ерекше қорғалатын табиғи аумақтарында *M. sieversii* түрінің генетикалық ресурстарын мониторингтеу мен сақтауға арналған сараланған іс-шараларды әзірлеудің кешенді ғылыми негізін құрайды.

Кілт сөздер: *Malus sieversii*; табиғи популяциялар; кеңістіктік таралу; биіктік талдауы; биоалуантүрлілікті мониторингтеу; ерекше қорғалатын табиғи аумақтар.

Пространственная организация, ДНК-баркодирование и создание банка ДНК диких популяций *Malus sieversii* для сохранения генетических ресурсов исчезающего вида

Исламова С.С.[1,2], Какимжанова А.А.[2], Дюсембекова Д.А.[2], Нургожанова Ж.С.[3], Садакова Б.Д.[4] Нуртаза А.С.[2]

НАО «Казахский агротехнический исследовательский университет имени С. Сейфуллина» Астана, Казахстан
ТОО «Национальный центр биотехнологии», Астана, Казахстан
Государственный национальный природный парк «Жонгар-Алатау», Сарканд, Казахстан
Государственный национальный природный парк «Тарбагатай», село Уржар, Казахстан

Аннотация

Предпосылки и цель. *Malus sieversii* (Ledeb.) M. Roem. – дикорастущий вид, являющийся основным предком культурных яблонь, занесённый в Красную книгу Республики Казахстан и признанный уязвимым видом по классификации МСОП. Сохранение природных популяций этого вида требует систематического мониторинга и детальной характеристики мест произрастания. В настоящей работе представлены результаты экспедиционных исследований на территории ГНПП «Жонгар-Алатау» и ГНПП «Тарбагатай», проведённых с целью описания пространственного распределения, высотных условий произрастания и молекулярно-генетической характеристики *M. sieversii.*

Материалы и методы. Обследовано 156 деревьев; для каждого образца зафиксированы GPS-координаты и высота над уровнем моря с использованием GPS. Пространственную структуру популяций оценивали на основе матрицы попарных географических расстояний с последующей геоинформационной обработкой и визуализацией в программе QGIS. Различия высотных диапазонов между участками анализировали с использованием непараметрического критерия

Краскела Уоллиса. Из листьев в трёх повторностях выделяли геномную ДНК, на основе которого был сформирован Банк ДНК-образцов. Видовую идентификацию подтверждали методом ДНК-баркодинга с использованием хлоропластных маркеров trnH-psbA, matK и rbcL.

Результаты. Пространственный анализ показал иерархическую пространственную структуру исследованных деревьев, включающую три уровня организации: локальный (внутри лесничеств), субрегиональный (между лесничествами Жонгар-Алатау) и региональный (обособление Тарбагатая). Высотный анализ выявил статистически значимые различия диапазонов произрастания между участками (тест Краскела-Уоллиса: H = 74,10; p<0,001): участки Жонгар-Алатау расположены на более высоких отметках (1091–1461 м, M ± SD = 1308 ± 104 м), Тарбагатай – к более низкому высотному поясу (834–1119 м, M ± SD = 888 ± 87 м). Для всех 156 образцов в трёх повторностях выделена ДНК и сформирован Банк ДНК-образцов; ДНК-баркодинг по маркёрам trnH-psbA, matK и rbcL подтвердил видовую принадлежность всех 11 проанализированных образцов (98–100%), полученные последовательности депонированы в GenBank с присвоением инвентарных номеров.

Заключение. Полученные результаты свидетельствуют о географической и экологической обособленности популяции Тарбагатая и в совокупности с созданным Банком ДНК формируют комплексную основу для разработки дифференцированных мер по мониторингу и сохранению генетических ресурсов *M. sieversii* на охраняемых территориях Казахстана.

Ключевые слова: *Malus sieversii*; дикорастущие популяции; пространственное распределение; высотный анализ; мониторинг биоразнообразия; охраняемые природные территории.