TÜBİTAK

Turkish Journal of Botany

http://journals.tubitak.gov.tr/botany/

Research Article

Turk J Bot
(2018) 42: 38-50

doi:10.3906/bot-1706-11

Plastid DNA variation of the endemic species *Oxytropis glandulosa* Turcz. (Fabaceae)

Alla KHOLINA[1,*], Marina KOZYRENKO[1], Elena ARTYUKOVA[1], Denis SANDANOV[2], Inessa SELYUTINA[3], Daba CHIMITOV[2]

[1]Federal Scientific Center of the East Asia Terrestrial Biodiversity, Far Eastern Branch of the Russian Academy of Sciences, Vladivostok, Russia

[2]Institute of General and Experimental Biology, Siberian Branch of the Russian Academy of Sciences, Ulan-Ude, Russia

[3]Central Siberian Botanical Garden, Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia

Received: 06.06.2017 • **Accepted/Published Online:** 15.09.2017 • **Final Version:** 11.01.2018

Abstract: *Oxytropis glandulosa* Turcz. (Fabaceae) is a rare perennial plant endemic to Buryatia (Russia). We sequenced three intergenic spacers (the *psbA–trnH*, *trnL–trnF*, and *trnS–trnG*) of chloroplast DNA (cpDNA) in specimens of four populations from two geographic regions (Barguzin and Yeravna depressions) within this species range. The levels of haplotype and nucleotide diversity varied in the range of 0.133–0.911 and 0.0002–0.0059, respectively. The highest values of these parameters are characteristic of populations located in the Barguzin depression. Variable sites detected within the intergenic spacers allowed the identification of eleven haplotypes; nine of them were found in the populations from the Barguzin depression and only two haplotypes were found in the populations from the Yeravna depression. No common haplotypes for all four populations were found. A high level of population differentiation (Φ_{ST} = 0.758, P < 0.0001) and the absence of evidence for isolation by distance were also found in *O. glandulosa*. Genealogical and phylogenetic analyses and Bayesian analysis of the genetic structure showed a clear division of the studied populations of *O. glandulosa* into three major lineages.

Key words: *Oxytropis glandulosa*, endemic, genetic diversity, intergenic spacer region, cpDNA

1. Introduction

As habitat fragmentation increases throughout the world it has many important ecological and genetic consequences for endemic plant populations. There are profound consequences for species genetic patterns as a result of this fragmentation and it is crucial to understand the characteristics of fragmented populations to provide adequate management, particularly where the populations and species are endangered. Knowledge of the population genetic structure of narrowly occurring endemic plants is of great importance for the purpose of conservation of the existing populations (Ellstrand and Elam, 1993). The analysis of variation in plant genomes with different modes of inheritance can provide valuable information on the genetic structure of current populations and on the evolutionary forces that have shaped the current distribution of genetic variation (Pleines et al., 2009 and references therein).

Oxytropis glandulosa Turcz. (Fabaceae) is a rare plant species endemic to Buryatia (region of Russia located in the south-central part of Siberia along the eastern shore of Lake Baikal). The species is included in the Red Data Book of the Russian Federation (Peshkova, 2008) and the Red Data Book of the Republic of Buryatia (Sandanov and Chimitov, 2013). *O. glandulosa* belongs to the section *Polyadena* Bunge (Malyshev, 2008). The species occurs in two isolated depressions more than 150 km distant from each other and separated by mountain ridges. The representatives of the populations from the Barguzin depression are diploids (Konichenko and Selyutina, 2013) and the representatives of the populations from the Yeravna depression are tetraploids (Zhukova, 1983). *O. glandulosa* habitats mostly are affected by plowing and grazing (Peshkova, 2008).

The chloroplast genome is useful in providing information on the inference of the evolutionary patterns and processes in plants (Raubeson and Jansen, 2005). At lower taxonomic levels, noncoding regions of the chloroplast DNA (cpDNA) are phylogenetically more informative than coding regions because they are under fewer functional constraints and evolve more rapidly (Gielly and Taberlet, 1994). Analysis of the chloroplast genome provides the resolution of controversial taxonomic and phylogenetic issues. The analysis of the cpDNA data reveals considerable divergence of closely related species of the genus *Oxytropis* DC., *O. chankaensis* Jurtz. and *O. oxyphylla* (Pall.) DC. (Artyukova and Kozyrenko, 2012). The phylogenetic relationships of the species and sections

* Correspondence: kholina@biosoil.ru

of the genus *Oxytropis* from Asian Russia were clarified on the basis of sequence comparison of the noncoding regions of cpDNA (Kholina et al., 2016).

In this study, we used sequence data for three noncoding regions (the *psbA–trnH*, *trnL–trnF*, and *trnS–trnG*) of the cpDNA to estimate the level of genetic diversity and population structure of *O. glandulosa.*

2. Material and methods

2.1. Population sampling

Narrow endemic *O. glandulosa* is restricted to areas within the Barguzin and Yeravna depressions. Currently only ten different localities of *O. glandulosa* are known, most of which are small and represented by a few individuals. Thus, the material for the study was collected only from four populations, which are large enough for sampling without damaging of the populations. Leaves from randomly selected plants approximately 100 m apart were collected. Sample size, population code, and the geographic coordinates for each population are given in Table 1. One accession of *O. triphylla* (Pall.) DC. (Russia, Buryatia, Barguzin depression, Kurumkansky District, near the village of Sahuli, 575 m a.s.l., 54.45°N, 110.45°E), from the section *Xerobia* Bunge, was used as an outgroup for phylogenetic analyses.

2.2. DNA extraction, amplification, and sequencing

Individual preparations of total DNA were extracted from the leaf tissues using the method described by Artyukova et al. (2004). We sequenced three intergenic spacer regions of the chloroplast genome: *psbA–trnH*, *trnL–trnF*, and *trnS–trnG*. The research was done using equipment of the Instrumental Centre of Biotechnology and Gene Engineering of FSCEATB FEB RAS. Double-stranded templates for direct sequencing were amplified on a thermal cycler UNO II 48 (Biometra, Germany) using the PCR parameters and universal primers, reaction conditions, and temperature regimes recommended for these regions (Taberlet et al., 1991; Shaw et al., 2005). The PCR products were sequenced using a BigDye terminator v. 3.1 sequencing standard kit (Applied Biosystems, USA). Sequencing was carried out in both directions under cyclic sequencing conditions and with the same pairs of primers as those used for amplification. In addition, internal primers were used for sequencing of the *trnL–trnF* and *trnS–trnG* regions (Taberlet et al., 1991; Shaw et al., 2005). The sequences were analysed on an ABI PRISM 3130 sequencer (Applied Biosystems, USA). Complete sequences were assembled using the Staden Package v. 1.5 (Bonfeld et al., 1995) and aligned manually with the program SeaView v. 4 (Gouy et al., 2010). The sequences were deposited in GenBank under accession numbers from LT732646 to LT732681.

2.3. Data analysis

For each individual, sequence data for the three regions were combined. The rate of plastid DNA evolution is very slow, and differences between substitution and indel mutation rates are unlikely to affect the resolution of intraspecific phylogenetic relationships (Gonzales et al., 2008). Each insertion or deletion, regardless of its size, was coded as a single mutational event in all analyses (Simmons and Ochoterena, 2000). The values of pairwise genetic distances (F_{ST}) among the populations, the number of haplotypes, the haplotype diversity (h), and nucleotide diversity (π) were calculated using the Arlequin v. 3.5 program (Excoffier and Lischer, 2010). To calculate the partitioning of genetic variation within and among the populations and/or population groups (fixation indices Φ_{ST}, Φ_{CT}, Φ_{SC}), we performed analyses of molecular variance (AMOVA; implemented in Arlequin). Gene flow (N_m) and the degree of divergence among the populations on the basis of nucleotide substitutions (K_S) were defined in the DnaSP v. 5.0 program (Librado and Rozas, 2009). We used the Mantel test with 10,000 permutations to analyze the relationships between the matrices of genetic (F_{ST}) and geographic distances using Arlequin. Phylogenetic analyses were carried out by the maximum likelihood (ML), neighbor joining (NJ), and maximum parsimony (MP) methods implemented in the PAUP v. 4.0b10 software package (Swofford, 2002). For the ML and MP analyses, heuristic searches were performed using tree-bisection-reconnection (TBR) branch swapping and 10 random sequence addition replicates. The GTR + I + G model was selected as the optimal setting for the ML and NJ analyses by Modeltest 3.6 (Posada and Crandall, 1998). The statistical significance of the branching order was assessed by bootstrap analysis with 1000 alternative trees (bootstrap percentage, BP). BP values below 50% were not considered. Two datasets were generated: a matrix containing data from the cpDNA regions of the 49 accessions of *O. glandulosa* and one accession of *O. triphylla*, and a matrix containing the sequences of *O. glandulosa* haplotypes G1–G11 and the sequences of *Oxytropis* species (haplotypes H1–H79 except H2, H19, H23, H29, H31, H39–H41, H45, H52, H55, H56) and the accessions of *Astragalus chinensis* L.f. (LM653160, LM653197, LM653234) and *A. davuricus* (Pall.) DC. (LM653161, LM653198, LM653235) from our previous study (Kholina et al., 2016). The genealogical relationships of haplotypes were determined using the median joining (MJ) method as implemented in the NETWORK v. 5 software program (Bandelt et al., 1999). The genetic structure of the total sample was examined for detecting the homogeneous groups within it. The analysis was performed with the BAPS6 software package (Corander et al., 2008), which uses a Bayesian approach and a stochastic optimization algorithm to assess the

Table 1. Sampling site locations, sample size, codes, and genetic diversity within populations of *Oxytropis glandulosa* (haplotype diversity and nucleotide diversity: mean value ± standard error).

The location of the population (no. of individuals)	Latitude, longitude	Altitude (m a.s.l.)	Code	Haplotypes	Haplotype frequency	Haplotype diversity (h)	Nucleotide diversity (π)
Russia, Buryatia, Barguzin depression, Kurumkansky District, near the village Argada (10)	54.37°N, 110.53°E	541	KUR	G1 G2 G3 G4 G5 G6 G7	0.100 0.200 0.100 0.100 0.100 0.300 0.100	0.911 ± 0.077	0.0059 ± 0.0033
Russia, Buryatia, Barguzin depression, Barguzin District, near the village Urzhil (14)	54.07°N, 110.39°E	501	BAR	G4 G6 G7 G8 G9	0.071 0.286 0.071 0.500 0.071	0.703 ± 0.101	0.0036 ± 0.0020
Russia, Buryatia, Yeravna depression, Yeravninsky District, near the village Shiringa (15)	52.67°N, 111.72°E	949	SHIR	G10 G11	0.933 0.067	0.133 ± 0.112	0.0002 ± 0.0002
Russia, Buryatia, Yeravna depression, Yeravninsky District, near the village Garam (10)	52.55°N, 111.48°E	947	GAR	G10 G11	0.800 0.200	0.356 ± 0.159	0.0006 ± 0.0004

likelihood of the individuals partitioning into a certain number of clusters (*K*). To determine the most probable number of clusters (without taking into account affiliation of the accessions to the populations), the "Clustering with linked loci" module of the program was utilized. The number of clusters for which the log of marginal likelihood is maximal corresponds to the most probable number of genetic groups in the sampled population. The analysis was repeated ten times for each value of *K*. In order to assess the probability of assigning the samples to the identified clusters and the possibility of mixed origin of the populations/samples, admixture analysis using the codon linkage model was performed.

3. Results

Nucleotide sequences of the *psbA–trnH*, *trnL–trnF*, and *trnS–trnG* intergenic spacers of cpDNA were obtained for 49 accessions of *O. glandulosa*. The variability of these regions was different. The lengths of the same regions in different accessions were different owing to the presence of insertions/deletions (indels), as well as mono- and dinucleotide repeats. Thus, the length of the *psbA–trnH* region ranged from 448 to 456 bp. The differences were determined by the presence of a poly-A motif, repeated 9 to 12 times, and a 6-bp indel that was a marker for population groups: the deletion was in the SHIR and GAR populations from the Yeravna depression and the insertion was in populations KUR and BAR from the Barguzin depression. No nucleotide substitutions in the sequences of this region were revealed. The length of the *trnL–trnF* spacer ranged from 747 to 772 bp. This variability was due to a mononucleotide repeat, a poly-T motif, repeated 8 to 11 times; a dinucleotide repeat, TA motif, repeated 8 to 13 times; and the 16-bp deletion in five accessions of the KUR and BAR populations. One nucleotide substitution was identified, which was informative for parsimony. The length of the *trnS–trnG* spacer ranged from 1185 to 1189 bp, the differences being due to 2 mononucleotide repeats, poly-A motif, repeated 10 to 14 times and poly-T motif, repeated 7 to 11 times. Within the *trnS–trnG*, six parsimony-informative nucleotide substitutions were identified, three of which were markers for population groups. The lengths of alignment matrices of the *trnH–psbA*, *trnL–trnF*, and *trnS–trnG* spacers were 456, 772, and 1192 bp, respectively. The total length of concatenated sequences of the three cpDNA regions was 2420 bp. The alignment contained 2363 monomorphic sites, 50 indels, and 7 variable sites that were parsimony-informative.

In the populations studied, the levels of haplotype and nucleotide diversity varied in the range of 0.133–0.911 and 0.0002–0.0059, respectively (Table 1). The highest values of these parameters are characteristic for the KUR and BAR populations, located in the Barguzin depression (the Barguzin group). The divergence of nucleotide sequences between KUR and BAR was low and absent among SHIR and GAR from the Yeravna depression (the Yeravna group) (Table 2). The highest sequence divergence was observed among the populations belonging to different

Table 2. Nucleotide divergence and genetic distances among the populations of *Oxytropis glandulosa*.

Population	KUR	BAR	SHIR	GAR
K_S				
KUR	–	2.000 (0)	5.000 (3)	5.000 (3)
BAR	0.00084	–	5.000 (5)	5.000 (5)
SHIR	0.00211	0.00211	–	0.000 (0)
GAR	0.00211	0.00211	0.00000	–
F_{ST}				
KUR	0.00000			
BAR	0.11071 ns	0.00000		
SHIR	0.81173*	0.86369*	0.00000	
GAR	0.75556*	0.82615*	–0.00264 ns	0.00000

K_S is nucleotide divergence: above the diagonal, average number of nucleotide differences between the populations (in parentheses, the number of fixed differences), below the diagonal, average number of nucleotide substitutions per site; F_{ST} is pairwise genetic distances; ns, not significant; * $P < 0.0001$ (1023 permutations). The population codes are given in Table 1.

groups. Average numbers of nucleotide differences and of nucleotide substitutions per site between two groups of populations were 5.000 (number of fixed differences: 3) and 0.00211, respectively. Considerable differentiation of the chloroplast genome among the populations growing in different geographically remote areas and therefore a high degree of genetic disunity between them are evidenced by the high pairwise genetic distances (F_{ST}, Table 2), as well as between the Barguzin and Yeravna groups (F_{ST} = 0.81038, P < 0.0001). At the same time, the genetic distances among the populations of the same geographic region were low and not significant (Table 2). The Mantel test showed that there was no significant pattern of isolation by distance in *O. glandulosa* because the matrix of geographic distances did not significantly correlate to F_{ST} (r = 0.973, P = 0.255). AMOVA revealed that the main part of total variation was due to variation among the populations (>75%) and less than 25% of the total genetic variance was within populations. Overall differentiation among all populations was high and significant (Φ_{ST} = 0.75783, P < 0.0001; Table 3), while the level of the inferred gene flow was low (N_m = 0.25 migrants per generation among populations). When the samples were divided into two groups according to their location in the Barguzin (KUR and BAR) and Yeravna (SHIR and GAR) depressions, the hierarchical analysis of genetic variation showed that the differences between the two groups accounted for 80% of total variance, while genetic differences among the populations within groups and intrapopulation variability accounted for 2.1% and 17.9%, respectively (Table 3). Gene flow between the KUR and BAR populations of the Barguzin group was 2.24 migrants per generation, and between the SHIR and GAR of the Yeravna group it was 5.94.

Nucleotide substitutions and indel variations revealed eleven different haplotypes (G1–G11). Of these, four haplotypes were found only in single specimens (private haplotypes). The KUR and BAR populations from the Barguzin depression contained nine haplotypes (G1–G9), while only two haplotypes (G10, G11) were found in the SHIR and GAR from the Yeravna depression. Three private haplotypes (G1, G3, G5) were found in KUR and only one private haplotype (G9) was found in BAR. In addition, common haplotypes were detected for populations from one group (KUR and BAR shared haplotypes G4, G6, and G7; SHIR and GAR shared haplotypes G10 and G11). No common haplotypes for all four populations were found. Haplotype frequencies are presented in Table 1. To demonstrate the genealogical relationships between haplotypes in *O. glandulosa*, the MJ network was constructed. All haplotypes were distributed in two divergent lineages/clades separated by more than fifteen mutation steps (Figure 1a). Clade A comprised the haplotypes found in populations SHIR and GAR from the Yeravna depression and clade B included all haplotypes found in the KUR and BAR populations from the Barguzin depression. The hypothetical haplotype (undetected in our

Table 3. Results of the analysis of molecular variance (AMOVA) in *Oxytropis glandulosa*.

Source of variation	Df	Sum of squares	Variance components	Variation (%)	Fixation index
One group: all the populations combined					
Among populations	3	339.448	9.10410	75.78	Φ_{ST} = 0.75783*
Within populations	45	130.919	2.90931	24.22	
Total	48	470.367	12.01341		
Two groups: populations from the Barguzin depression versus populations from the Yeravna depression					
Among regions	1	325.557	12.99847	80.00	Φ_{CT} = 0.79996 ns
Among populations within regions	2	13.891	0.34108	2.10	Φ_{SC} = 0.10494 ns
Within populations	45	130.919	2.90931	17.90	Φ_{ST} = 0.82095*
Total	48	470.367	16.24887		
One group defined according BAPS clusters					
Among populations	2	363.197	12.63908	84.44	Φ_{ST} = 0.84436*
Within populations	46	107.171	2.32979	15.56	
Total	48	470.367	14.96888		

Df is degrees of freedom; Φ_{ST} is correlation within populations relative to the total; Φ_{CT} is correlation of individuals within groups relative to the total; Φ_{SC} is correlation within populations relative to groups; ns, not significant; * P < 0.0001 (1023 permutations).

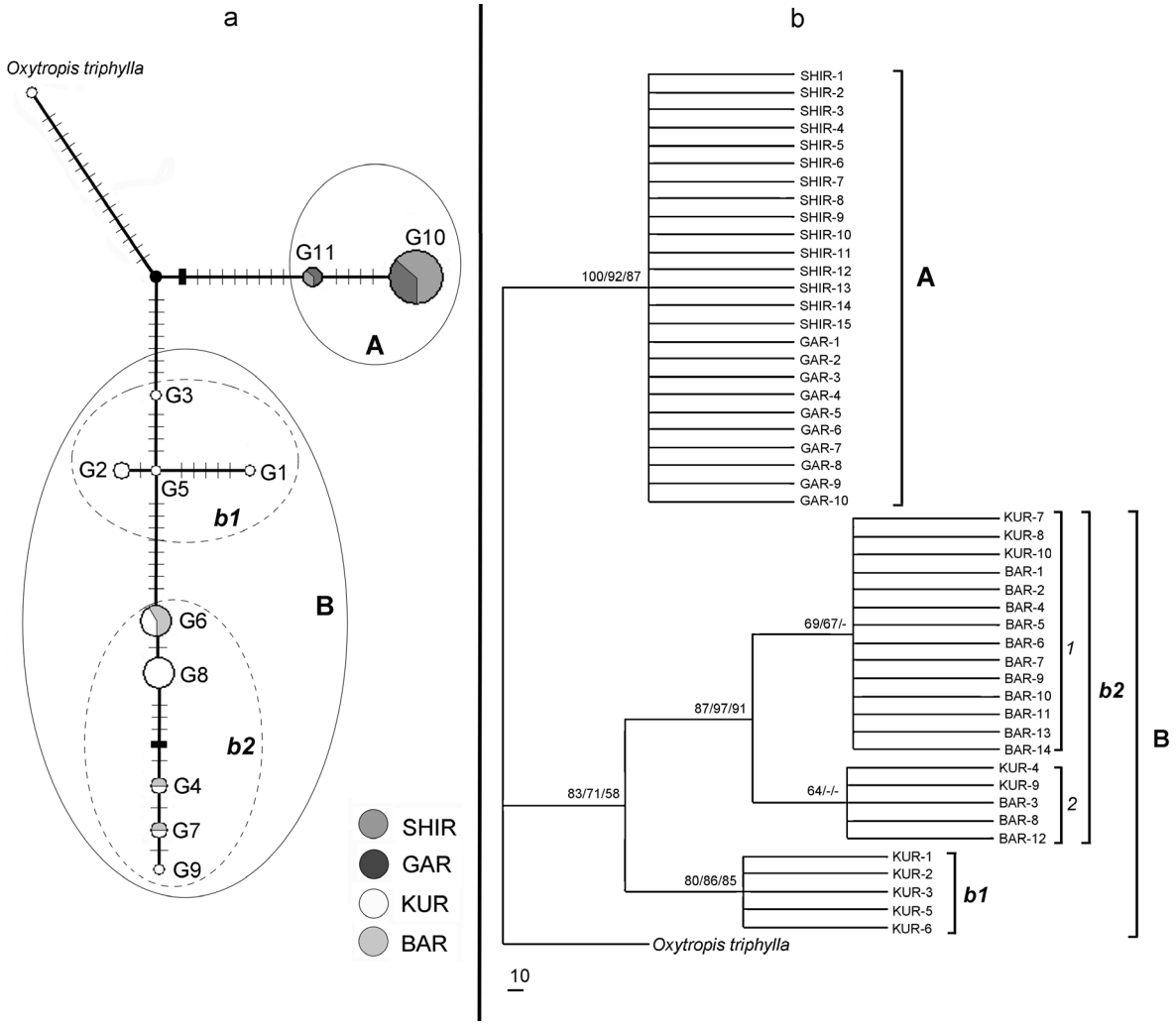

Figure 1. Genealogical relationships between the representatives of four populations of *Oxytropis glandulosa*. a) Network of haplotypes constructed with the help of the MJ method. Opened circles represent sampled haplotypes of *O. glandulosa* (G1–G11); their relative sizes are proportional to haplotype frequencies. Black dots represent hypothetical (missing) haplotype; thin black bars depict nucleotide substitutions; thick black bars depict multibase indels. Haplotype lineages/clades A and B are surrounded by a continuous line; the two subclades are encircled with dashed line. b) MP tree (tree length of 48 steps, CI = 0.6667, HI = 0.3333, RI = 0.9429). Numbers at the node designate the bootstrap values calculated for the MP/NJ/ML analyses (above 50%). The scale bar below corresponds to the branch length unit of 10 steps. The sequences of *O. triphylla* were used as outgroup. The population codes: KUR – Barguzin depression, Kurumkansky District; BAR – Barguzin depression, Barguzin District; SHIR – Yeravna depression, Yeravninsky District, near the village Shiringa; GAR – Yeravna depression, Yeravninsky District, near the village Garam.

study or extinct) occupies an intermediate position among these lineages. In addition, in clade B two subclades (*b1* and *b2*, Figure 1a) can be identified; the first subclade comprised only haplotypes of the KUR population (G1, G2, G3, and G5) and the second one was formed by other haplotypes of the KUR and all haplotypes of the BAR population.

The phylogenetic reconstruction of genetic relationships was based on the dataset of 49 *O. glandulosa* samples and one accession of *O. triphylla* consisting of 2426 characters. There were ten variable sites, of which seven sites were parsimony-informative. The phylogenetic trees constructed using different methods (MP, NJ, ML) do not differ in topology. The MP consensus tree (Figure 1b) was consistent with the topology of the haplotype network generated in the NETWORK program (Figure 1a). All accessions of populations SHIR and GAR (Yeravna depression) with high statistical support (BP 100%, 92%, 87% in MP, NJ, ML analyses, respectively) were included in clade A, and all accessions of the populations KUR and

BAR (Barguzin depression) were included in clade B (BP 83%, 71%, 58%); these clades remained unresolved. Clades A and B differed by the presence of a 6-bp indel in the *psbA–trnH* (positions 162 to 167) spacer and synapomorphic substitutions in the *trnL–trnF* (A↔C, position 1290) and *trnS–trnG* (A↔T, position 1398; C↔A, position 2194) spacers. The accessions of clade B were arranged in two highly supported subclades, *b1* and *b2* (Figure 1b). Five accessions (KUR1, KUR2, KUR3, KUR5, and KUR6) of subclade *b1* were resolved as a distinct group (Figure 1b) characterized by the synapomorphic substitutions in the *trnL–trnF* (A, position 601) and *trnS–trnG* (T, position 1580; C, positions 1722 and 2108) spacers. Subclade *b2* was split into two lowly supported groups; one of them comprised five representatives of two populations (KUR4, KUR9, BAR3, BAR8, and BAR12). The sequences of these samples differed from the other samples of subclade *b2* in the synapomorphic deletion of 16 bp in the *trnL–trnF* (positions 769 to 784) spacer. Thus, analysis revealed 4 divergent groups that did not correspond to the populations studied. The divergence of the nucleotide sequences between 4 groups is shown in Table 4. The high sequence divergence was observed between all groups except for two groups constituting the subclade *b2* (Figure 1b). Genetic distances between all groups were high. However, the lack of nucleotide divergence and differentiating nucleotide substitutions among two groups in highly supported subclade *b2* (Table 4) indicates the close relationships in the recent past between populations KUR and BAR, which are currently geographically disconnected.

Bayesian analysis of the genetic structure of *O. glandulosa* samples regardless of their population assignment showed (P = 0.66113) subdivision into 3 clusters (Figure 2a). Cluster 1 was formed by all accessions of the SHIR and GAR populations (Yeravna depression). Most individuals of the KUR and BAR populations (Barguzin depression) were found in cluster 2. Cluster 3 consisted of the five individuals from the KUR population (Figure 2a). Analysis of the genetic composition of haplotypes showed a statistically significant genetic homogeneity of populations SHIR and GAR (cluster 1) and five samples from the KUR population (cluster 3). In contrast to these clusters, three individuals (KUR9, BAR3, and BAR8) in cluster 2 were genetically heterogeneous (Figure 2b). These haplotypes contained small and equal proportions of the mutations specific to cluster 3, pointing to their mixed origin. AMOVA revealed that the differentiation among the clusters was high and significant (Φ_{ST} = 0.84436, P <

Table 4. Nucleotide divergence and genetic distances between the four groups identified in the phylogenetic analysis of 49 accessions *Oxytropis glandulosa*.

Group	1	2	3	4
K_S				
1	–	6.000 (6)	5.000 (5)	5.000 (5)
2	0.00251	–	0.000 (0)	5.000 (5)
3	0.00210	0.00000	–	4.000 (4)
4	0.00209	0.00209	0.00168	–
F_{ST}				
1	0.00000			
2	0.97472*	0.00000		
3	0.97927*	0.97188*	0.00000	
4	0.95199*	0.85972*	0.91391**	0.00000

K_S is nucleotide divergence: above the diagonal, average number of nucleotide differences among the populations (in parentheses, the number of fixed differences); below the diagonal, average number of nucleotide substitutions per site; F_{ST} is pairwise genetic distances. Groups were defined according to the results of phylogenetic analysis (Figure 2b): 1, clade A containing all samples of the SHIR and GAR populations; 2, group 1 of subclade *b2* containing fourteen samples from the populations KUR and BAR; 3, group 2 of subclade *b2* containing five samples from the populations KUR and BAR; 4, subclade *b1* containing five samples from the KUR population. * P < 0.0001, ** P < 0.01 (1023 permutations). The population codes are given in Table 1.

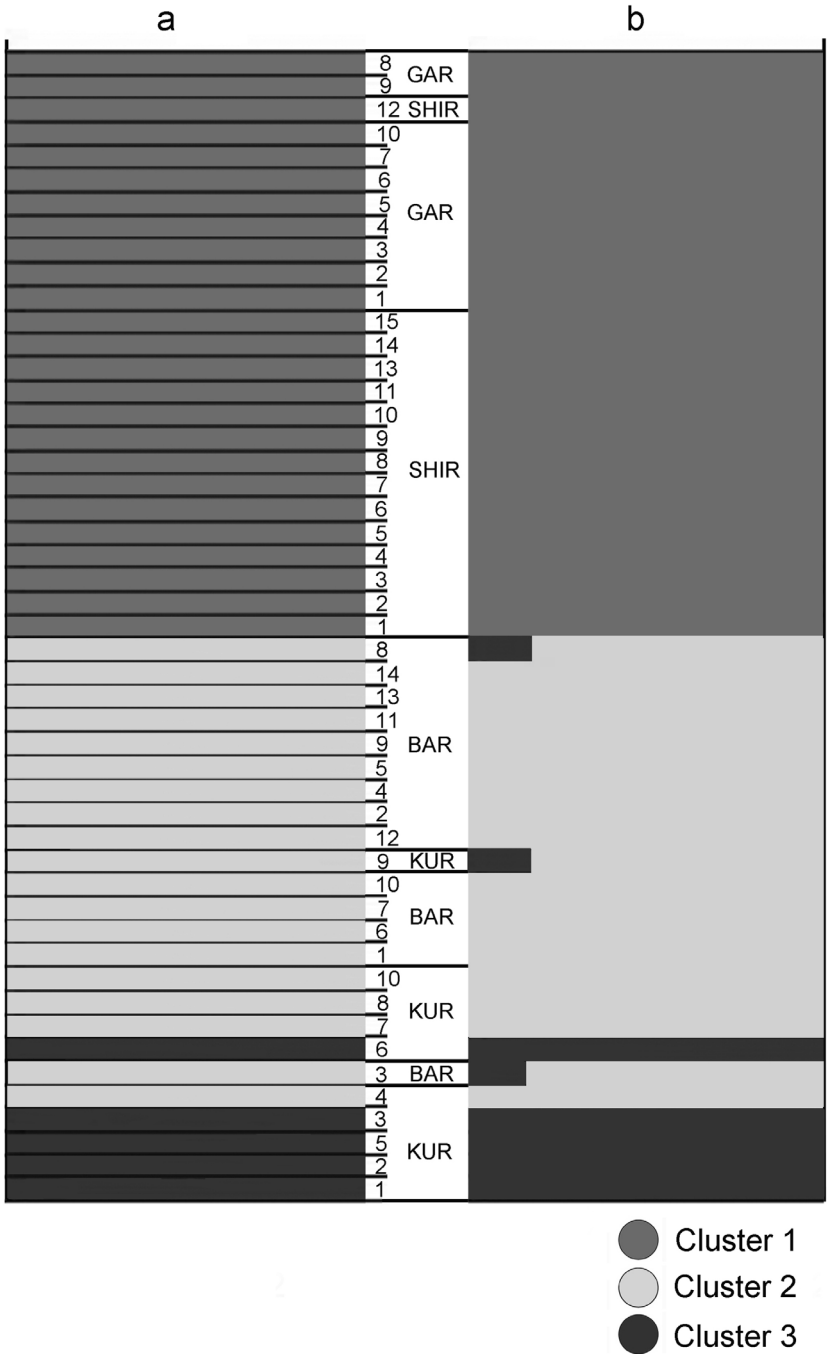

Figure 2. Bayesian analysis of the genetic structure of *Oxytropis glandulosa*: a) clustering of 49 accessions regardless of their population assignment; b) probability of the accession assignment to one of the three clusters in the admixture analysis. Each individual is represented with a horizontal bar and each inferred cluster is marked with a different gray tone. Columns are shaded with different colors in proportions corresponding to estimated admixture coefficients for each individual. Individuals from the different populations are separated by black horizontal lines. The population codes are given in Table 1.

0.0001) and similar to that among the four populations (Table 3).

The sequences of eleven haplotypes of *O. glandulosa* and 67 accessions representing 59 species of the genus *Oxytropis*, as well as members of the genus *Astragalus* (as outgroup taxa), obtained earlier (Kholina et al., 2016) were used for revealing phylogenetic relationships of *O. glandulosa* with other members of the genus. The dataset

of 80 accessions contained 2961 characters. There were 131 variable sites, of which 74 sites were parsimony-informative. The overall topology of the phylogenetic tree generated from this dataset (Figure 3) did not differ from that obtained previously (Kholina et al., 2016). The *Oxytropis* species formed two clades with high statistical support (Figure 3). All haplotypes of *O. glandulosa* were placed in clade II (BP 100%, 99%, 100% in MP, NJ, and ML analyses, respectively). Phylogenetic relationships between most taxa in clade II remain unsolved. The cause of this difficult species delimitation may be related to recent speciation, hybridization, introgression events, and the presence of intraspecific polyploid series, as in the genus *Onobrychis* Mill. (Lewke Bandara et al., 2013). Statistically supported groups, formed by the species of one section or by the species from different sections and by the haplotypes of *O. glandulosa*, can be identified within clade II. The haplotypes of *O. glandulosa* were divided into three groups. One group with low support (BP 67%, 65%, 61%) contained haplotypes G1, G2, G3, and G5 of the KUR population; the second group with moderate support (BP 76%, 89%, 57%) was formed by haplotypes G4, G6, G7, G8, and G9 of populations KUR and BAR; and the third group contained haplotypes G10 and G11 of populations SHIR and GAR (BP 88%, 52% in MP and NJ analyses, respectively).

Thus, the genealogical and phylogenetic reconstructions of relationships and Bayesian analysis of the genetic structure showed a clear subdivision of the pattern haplotypes into three major cpDNA lineages in *O. glandulosa*.

4. Discussion

Many rare and endemic species exist as small isolated populations, whose fragmentation increases under anthropogenic pressure. As a rule, the species with narrow ranges have low genetic polymorphism (Ellstrand and Elam, 1993). Our data on sequence variability of the three cpDNA regions in *O. glandulosa* showed that the populations from the Barguzin depression are characterized by a high level of haplotype diversity (0.911 and 0.703), while a low level of haplotype diversity (0.133 and 0.356) was detected in the populations from the Yeravna depression (Table 1). The high level of haplotype diversity in the Barguzin populations is similar to those of other narrow species from the family Fabaceae, such as endemic *O. chankaensis* (0.718, Artyukova et al., 2011), *Hymenaea stigonocarpa* Mart. ex Hayne (0.804; Ramos et al., 2007), and *Dalbergia nigra* Allem. ex Benth. (0.752; Ribeiro et al., 2011). Low levels of nucleotide diversity in the Yeravna populations could be due to several factors. It may originate from a founder population containing only a small fraction of the genetic variation present in its progenitor species, which may have experienced a bottleneck effect in its evolutionary history. Also, isolated populations of this region were apt to suffer from genetic drift that may have contributed to the lack of genetic diversity. The low level of nucleotide diversity in *O. glandulosa* (range: 0.0002–0.0059; Table 1) within a total DNA region of 2420 bp is consistent with the low mutation rate in the chloroplast genome that has been estimated for the genus *Oxytropis* of 8.9×10^{-10} substitutions per site per year (Wojciechowski, 2005). Likewise, a low level of nucleotide diversity in cpDNA has been found in endemic legume species such as *O. chankaensis* (0.0005, Artyukova et al., 2011), *Hymenaea stigonocarpa* (range: 0–0.0027, Ramos et al., 2007), and *Dalbergia nigra* (range: 0–0.00084, Ribeiro et al., 2011). The high differentiation of populations in *O. glandulosa* ($\Phi_{ST} = 0.758$; Table 2) is similar to that in *Primula secundiflora* Franch. (0.816; Wang et al., 2008). A high level of population differentiation and the absence of shared haplotypes between two groups of *O. glandulosa* populations from different geographical regions (Tables 1 and 3) indicated that gene flow was limited.

All studied plants were distributed in two major divergent clades (A and B, Figure 1) according to their geographical locations and levels of ploidy. Clade B was split into two subclades, *b1* and *b2*. Thus, three cpDNA lineages (clade/subclade in our analyses) were identified in populations of *O. glandulosa*: lineage A comprised two haplotypes found in the Yeravna populations, lineage *b1* included the only four haplotypes from the KUR population of the Barguzin depression, and lineage *b2* included the majority of haplotypes of the two Barguzin populations. Lineages A and *b1* were characterized by the synapomorphic substitutions in the *trnL–trnF* and *trnS–trnG* spacers. The genetic divergence revealed between three lineages (Table 4) was high and comparable to the divergence among the *Oxytropis* species within different sections (Artyukova and Kozyrenko, 2012; Kholina at al., 2016) and between closely related species of other genera, for instance *Megadenia* (Artyukova et al., 2014) and *Iris* series *Psammiris* (Kozyrenko et al., 2009).

The genetically similar populations from the Barguzin depression and populations from the Yeravna depression are geographically distant from each other. In fact, the distributional areas of *O. glandulosa* are isolated and fragmented by mountains and valleys, which probably impose significant barriers to gene exchange. Thus, the present genetic structure of *O. glandulosa* was mainly shaped by the fragmentation of ancestral populations. However, we did not find any correlation between genetic and geographic distances using the Mantel test. In addition, the presence of synapomorphic sites in each group, the absence of common haplotypes, the placement in the two divergent groups in the phylogenetic and genealogical

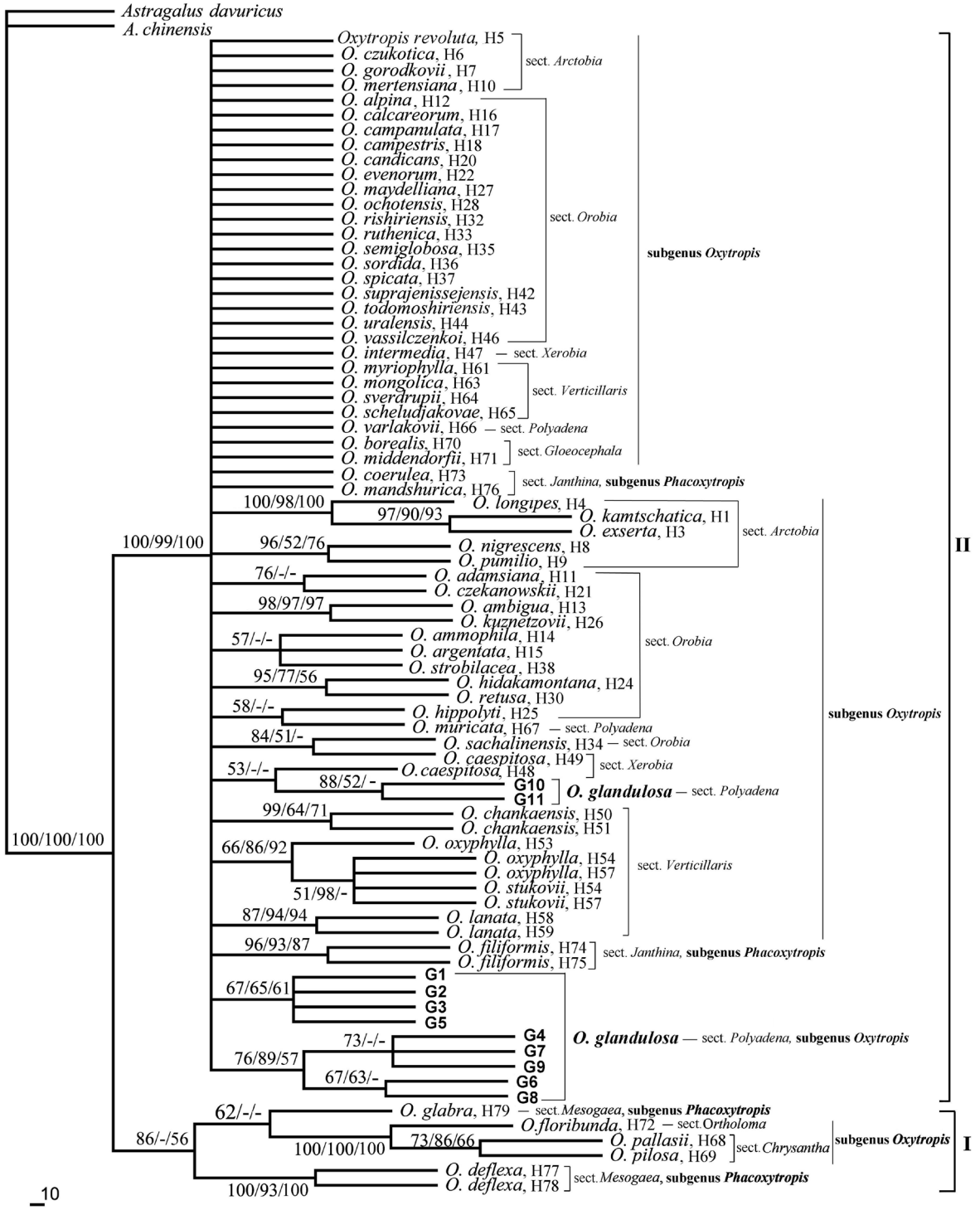

Figure 3. MP tree of the representatives of the genus *Oxytropis* (tree length of 559 steps; CI = 0.6941; HI = 0.3059; RI = 0.8025). Numbers at the nodes designate the bootstrap support values calculated for the MP/NJ/ML analyses (above 50%). H1 to H79, haplotypes of *Oxytropis* species (Kholina et al., 2016); G1 to G11, haplotypes of *O. glandulosa*. The sequences of *Astragalus davuricus* and *A. chinensis* were used as an outgroup. The names of the sections and subgenera are given according to Malyshev (2008). The scale bar below corresponds to the branch length unit of 10 steps.

constructions, and the lack of gene flow (via seed) suggest that *Oxytropis* populations from the Barguzin and Yeravna depressions probably have their own independent evolutionary history.

Polyploidy is common in the genus *Oxytropis* (Malyshev, 2008; Martin et al., 2015). Some species have a variable number of chromosomes (2n): 16 and 32; 16, 32, and 48; 48 and 64; and 48, 64, and 96. This may be an appearance of racial differentiation in species with broad habitat or when the species inhabit preferably in several altitudinal belts. Levels of ploidy of *O. glandulosa* from the Barguzin and Yeravna depressions are different (Zhukova, 1983; Konichenko and Selyutina, 2013). Reduced genetic variation in the tetraploid populations from the Yeravna depression is probably caused by founder effects during the polyploid speciation, as was shown in polyploid plants (Guo et al., 2013). The considerable genetic divergence that we observed between two groups of *O. glandulosa* populations most likely has resulted from chromosomal isolation (i.e. their hybrids presumably would be triploid and therefore sterile).

The deep split between three lineages of *O. glandulosa* is suggestive of cryptic diversity. Cryptic species are defined as two or more distinct lineages that are classified as a single nominal species, owing to the fact that they are difficult or impossible to be distinguished based on morphological characters alone (Ma et al., 2015; Shneyer and Kotseruba, 2015). The existence of cryptic species in plants was associated primarily with polyploidy. It is widely acknowledged that polyploidy has contributed greatly to plant diversification (Soltis et al., 2007, 2010; Wood et al., 2009). More than 15% of angiosperm species have been determined to be polyploids (Wood et al., 2009). In addition, polyploids have also been found in numerous diploid-polyploid morphological species (Soltis et al., 2010). Due to reproductive isolation and termination of gene flow among diploids and polyploids, these could have developed into different biological species under morphological stasis. A group of competent authors proposed to describe autopolyploid cytotypes as separate species (Soltis et al., 2007). Cryptic species produced by polyploidy were revealed in some species or species complexes (Liu et al., 2013; Britton et al., 2014; Ma et al., 2015), including legume species such as *Leucaena* species (Govindaralulu et al., 2011) or *Medicago prostrata* Jacq. (Eriksson et al., 2017). A group of individual organisms on an evolutionary trajectory independent from all other groups can be regarded as a separate species, i.e. they are separately evolving lineages, sensu De Queiroz (2007). Some species studied (Ma et al., 2015; Eriksson et al., 2017) comprise multiple, evolutionarily independent lineages, which, despite their potentially cryptic nature, qualify as species under the unified species concept (De Queiroz, 2007). Presumably tetraploid *O. glandulosa* populations should be regarded as a different species from the diploids, despite no known morphological distinctiveness between them.

Environmental conditions, landscape complexity, orogeny, and glaciation are considered as other drivers of cryptic speciation. For instance, for *Caesalpinia hintonii* complex, the volcanic activity, and geological heterogeneity of the distribution area have had considerable influence on population divergence (Sotuyo et al., 2007); allopatric divergence between two legume species of *Ammopiptanthus* correlated with geographic barriers and climate changes during glacial periods (Su et al., 2016). The differences in altitude of the habitats and absence of gene flow caused the appearance of two separate species in *Pedicularis chamissonis* Steven (Fujii et al., 2001). In *Tetraria triangularis* (Boeck.) C.B.Clarke, cryptic species formation occurs due to complex topography and topological barriers to gene flow (Britton et al., 2014), while in *Taxus wallichiana* Zucc. cryptic speciation is associated with the combination of both geological and climatic events (Liu et al., 2013). The formal recognition of cryptic species is desirable not only because it better represents the natural order but also because it has implications for biological diversity conservation (Bickford et al., 2006; Shneyer and Kotseruba, 2015).

The pattern of the cpDNA haplotype distribution of *O. glandulosa* may reflect the presence of several cryptic *Oxytropis* species vicariously replacing *O. glandulosa* or coexisting with it in Buryatia. This question requires a more detailed analysis of genetic and morphological data of extended sampling of plants from Buryatia and Mongolia.

Acknowledgments

This work was supported by the Russian Foundation for Basic Research (grant no. 16-04-01399) and the Program of the Presidium of the RAS (project no. 15-1-6-030) and was partly supported by a project of the Russian Federal Budget (VI.52.1.7).

References

Artyukova EV, Kholina AB, Kozyrenko MM, Zhuravlev YuN (2004). Analysis of genetic variation in rare endemic species *Oxytropis chankaensis* Jurtz. (Fabaceae) using RAPD markers. Russ J Genet 40: 710-716.

Artyukova EV, Kozyrenko MM (2012). Phylogenetic relationships of *Oxytropis chankaensis* Jurtz. and *Oxytropis oxyphylla* (Pall.) DC. (Fabaceae) inferred from the data of sequencing of the ITS region of the nuclear ribosomal DNA operon and intergenic spacers of the chloroplast genome. Russ J Genet 48: 163-169.

Artyukova EV, Kozyrenko MM, Boltenkov EV, Gorovoy PG (2014). One or three species in *Megadenia* (Brassicaceae): insight from molecular studies. Genetica 142: 337-350.

Artyukova EV, Kozyrenko MM, Kholina AB, Zhuravlev YuN (2011). High chloroplast haplotype diversity in the endemic legume *Oxytropis chankaensis* may result from independent polyploidization events. Genetica 139: 221-232.

Bandelt HJ, Forster P, Röhl A (1999). Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol 16: 37-48.

Bickford D, Lohman DJ, Sodhi NS, Ng PKL, Meier R, Winker K, Ingram KK, Das I (2006). Cryptic species as a window on diversity and conservation. Trends Ecol Evol 22: 148-155.

Bonfeld JK, Smith KF, Staden R (1995). A new DNA sequence assembly program. Nucleic Acids Res 23: 4992-4999.

Britton MN, Hedderson TA, Verboom GA (2014). Topography as a driver of cryptic speciation in the high elevation cape sedge *Tetraria triangularis* (Boeck.) C. B. Clarke (Cyperaceae: Schoeneae). Mol Phylogenet Evol 77: 96-109.

Corander J, Marttinen P, Sirén J, Tang J (2008). Enhanced Bayesian modelling in BAPS software for learning genetic structures of populations. BMC Bioinformatics 9: 539.

De Queiroz K (2007). Species concepts and species delimitation. Syst Biol 56: 879-886.

Doyle JJ, Doyle JL, Rauscher JT, Brown AHD (2004). Diploid and polyploid reticulate evolution throughout the history of the perennial soybeans (*Glycine* subgenus *Glycine*). New Phytol 161: 121-132.

Ellstrand NC, Elam DR (1993). Population genetic consequences of small population size: implication for plant conservation. Annu Rev Ecol Syst 24: 217-242.

Eriksson JS, Blanco-Pastor JL., Sousa F, Bertrand YJK, Pfeil BE (2017). A cryptic species produced by autopolyploidy and subsequent introgression involving *Medicago prostrata* (Fabaceae). Mol Phylogenet Evol 107: 367-381.

Excoffier L, Lischer HEL (2010). Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour 10: 564-567.

Fujii N, Ueda K, Watano Y, Shimizu T (2001). Two genotypes of *Pedicularis chamissonis* (Scrophulariaceae) distributed at Mt. Gassan, Japan: additional genetic and morphological studies. J Plant Res 114: 133-140.

Gielly L, Taberlet P (1994). The use of chloroplast DNA to resolve plant phylogenies: noncoding versus *rbcL* sequences. Mol Biol Evol 11: 769-777.

Gonzales E, Hamrick JL, Chang SM (2008). Identification of glacial refugia in south-eastern North America by phylogeographical analyses of a forest understorey plant, *Trillium cuneatum*. J Biogeogr 35: 844-852.

Gouy M, Guindon S, Gascuel O (2010). SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Mol Biol Evol 27: 221-224.

Govindaralulu R, Hughes CE, Bailey CD (2011). Phylogenetic and population genetic analyses of diploid *Leucaena* (Leguminosae; Mimosoideae) reveal cryptic species diversity and patterns of divergent allopatric speciation. Am J Bot 98: 2049-2063.

Guo YP, Tong XY, Wang LW, Vogl C (2013). A population genetic model to infer allotetraploid speciation and long-term evolution applied to two yarrow species. New Phytol 199: 609-621.

Kholina AB, Kozyrenko MM, Artyukova EV, Sandanov DV, Andrianova EA (2016). Phylogenetic Relationships of the species of *Oxytropis* DC. subg. *Oxytropis* and *Phacoxytropis* (Fabaceae) from Asian Russia inferred from the nucleotide sequence analysis of the intergenic spacers of the chloroplast genome. Russ J Genet 52: 780-793.

Konichenko ES, Selyutina IYu (2013). Chromosome numbers of rare and endemic species of the genus *Oxytropis* (Fabaceae). Botanicheskii Zhurnal 98: 647-651 (in Russian with abstract in English).

Kozyrenko MM, Artyukova EV, Zhuravlev YuN (2009). Independent species status of *Iris vorobievii* N.S. Pavlova, *Iris mandshurica* Maxim., and *Iris humilis* Georgi (Iridaceae): Evidence from the nuclear and chloroplast genomes. Russ J Genet 45: 1394-1402.

Lewke Bandara N, Papini A, Mosti S, Brown T, Smith LMJ (2013). A phylogenetic analysis of genus *Onobrychis* and its relationships within the tribe *Hedysareae* (Fabaceae). Turk J Bot 37: 981-992.

Librado P, Rozas J (2009). DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25: 1451-1452.

Liu J, Moller M, Provan J, Gao LM, Poudel RC, Li DZ (2013). Geological and ecological factors drive cryptic speciation of yews in a biodiversity hotspot. New Phytol 199: 1093-1108.

Ma XG, Zhao C, Wang CB, Liang QL, He XJ (2015). Phylogenetic analyses and chromosome counts reveal multiple cryptic species in *Bupleurum commelynoideum* (Apiaceae). J Syst Evol 53: 104-116.

Malyshev LI (2008). Diversity of the genus *Oxytropis* in the Asian part of Russia. Turczaninowia 11: 5-141 (in Russian with abstract in English).

Martin E, Karaman Erkul S, Aytaç Z (2015). Karyological studies on *Oxytropis* (Fabaceae) from Turkey. Caryologia 68: 357-362.

Peshkova GA (2008). Glandulous locoweed – *Oxytropis glandulosa* Turcz. In: Trutnev YuP, editor. Red Data Book of Russian Federation (Plants and Fungi). Moscow, Russia: KMK Scientific Press, p. 253 (in Russian).

Pleines T, Jakob SS, Blattner FR (2009). Application of non-coding DNA regions in intraspecific analyses. Plant Syst Evol 282: 281-294.

Posada D, Crandall KA (1998). Modeltest: testing the model of DNA substitution. Bioinformatics 14: 817-818.

Ramos ACS, Lemos-Filho JP, Ribeiro RA, Santos FCR, Lovato MB (2007). Phylogeography of the tree *Hymenaea stigonocarpa* (Fabaceae: Caesalpinioideae) and the influence of quaternary climate changes in the Brazilian Cerrado. Ann Bot-London 100: 1219-1228.

Raubeson LA, Jansen RK (2005). Chloroplast genomes of plants. In: Henry RJ, editor. Plant Diversity and Evolution: Genotypic and Phenotypic Variation in Higher Plants. Cambridge, MA, USA: CABI, pp. 45-68.

Ribeiro RA, Lemos-Filho JP, Ramos ACS, Lovato MB (2011). Phylogeography of the endangered rosewood *Dalbergia nigra* (Fabaceae): insights into the evolutionary history and conservation of the Brazilian Atlantic Forest. Heredity 106: 46-57.

Sandanov DV, Chimitov DG (2013). Glandulous locoweed – *Oxytropis glandulosa* Turcz. In: Pronin NM, editor. Red Data Book of Republic of Buryatia: Rare and Endangered Species of Animals, Plants and Fungi. 3rd ed. Ulan-Ude, Russia: Buryat Scientific Center SB RAS Publisher, p. 526 (in Russian).

Shaw J, Lickey EB, Beck JT, Farmer SB, Liu W, Miller J, Siripun KC, Winder CT, Schilling EE, Small RL (2005). The tortoise and the hare II: Relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. Am J Bot 92: 142-166.

Shneyer VS, Kotseruba VV (2015). Cryptic species in plants and their detection by genetic differentiation between populations. Russ J Genet Appl Res 5: 528-541.

Simmons MP, Ochoterena H (2000). Gaps as characters in sequence-based phylogenetic analyses. Syst Biol 49: 369-381.

Soltis DE, Buggs RJA, Doyle JJ, Soltis PS (2010). What we still don't know about polyploidy. Taxon 59: 1387-1403.

Soltis DE, Soltis PS, Schemske DW, Hancock JF, Thompson JN, Husband BC, Walter S, Judd WS (2007). Autopolyploidy in angiosperms: have we grossly underestimated the number of species? Taxon 56: 13-30.

Sotuyo S, Delgado-Salinas A, Chase MW, Lewis JP, Oyama K (2007). Cryptic speciation in the *Caesalpinia hintonii* complex (Leguminosae: Caesalpinioideae) in a seasonally dry Mexican Forest. Ann Bot-London 100: 1307-1314.

Su Z, Pan B, Zhang M, Shi W (2016). Conservation genetics and geographic patterns of genetic variation of endangered shrub *Ammopiptanthus* (Fabaceae) in northwestern China. Conserv Genet 17: 485-496.

Swofford DL (2002). PAUP* Phylogenetic Analysis Using Parsimony (and Other Methods), Version 4.0b10. Sunderland, MA, USA: Sinauer Associates, Inc.

Taberlet P, Gielly L, Pautou G, Bouvet J (1991). Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Mol Biol 17: 1105-1109.

Wang FY, Gong X, Hu CM, Hao G (2008). Phylogeography of an alpine species *Primula secundiflora* inferred from the chloroplast DNA sequence variation. J Syst Evol 46: 13-22.

Wojciechowski MF (2005). Astragalus (Fabaceae): a molecular phylogenetic perspective. Brittonia 57: 382-396.

Wood TE, Takebayashi N, Barker MS, Mayrose I, Greenspoon PB, Riesebergb LH (2009). The frequency of polyploid speciation in vascular plants. P Natl Acad Sci USA 106: 13875-13879.

Zhukova PG (1983). Chromosome numbers of some species of the Family Fabaceae from North-East Asia. Botanicheskii Zhurnal 68: 925-932 (in Russian with abstract in English).