

TrnL-trnF cpDNA polymorphism in some representatives of the genus *Betula*

Svetlana Medvedeva¹, Olga Cherepanova¹, Olga Tolkach¹, Vasily Ponomarev¹, Galina Malosieva²

¹ Botanical Garden, UrB RAS, 202 a, 8 March, Ekaterinburg, 620144, Russia

² Botanical Garden named after E. Gareeva of National Academy of Sciences, 1a, I. Akhunbaeva str., Bishkek, 720064

Abstract. We present preliminary results of the trnL-trnF cpDNA region analysis for some representatives of the g. *Betula*, out of which *B. procurva* can be considered the most interesting. The disjunctive range of this species is confined to the mountain system in southeast Central Asia (Pamir-Alai). The occurrence of the birch is isolated from the main range in the mountainous part of the Urals, in the boreal forests zone and in the Trans-Ural forest-steppe. Due to high interspecific hybridization and population variability of the g. *Betula*, doubts were raised about the correct identification of the representatives of *B. procurva*. The studied genetic variability and population structure of *B. procurva*, *B. nana*, *B. pubescens*, *B. turkestanica*, *B. tianshanica* and *B. pendula* species indicate active introgression and hybridization processes. Polymorphism in all groups is significantly reduced, increasing slightly when comparing the most distant groups. The matching of the *B. procurva* cpDNA haplotypes (GBS) with *B. pendula*, *B. nana*, and *B. pubescens* shows that this trnL-trnF cpDNA site cannot be used for molecular identification of birch species by barcoding as a single marker, but this marker use is possible for determination of certain *B. procurva* population. Based on the studied cpDNA region (trnL-F) we found a clear geographic subdivision in *B. procurva* populations of the Trans-Urals and Central Asia.

1 Introduction

Conservation and reproduction of forests is one of the priority areas of any country's environmental policy. The forest fund of Russia plays an important role in the economic sector, ensuring the performance of various national economy sectors: building, medicine, food industry, etc.

Changes in climatic conditions associated with increased fires, uncontrolled logging, and increased technogenic pollution lead to reduction in the area of natural forests as well as natural regeneration of forests and the subsequent transformation of established ecosystems. In order to protect the gene pool of the main forest tree species, genetic reserves are being created and the work is being done to conserve the existing genetic diversity. In line with abovementioned the study of polymorphism, phylogeography, and phylogeny of the main forest tree species populations and birches in particular is a relevant field of genetic research of the forest fund.

As far as fundamental science is concerned, genetic diversity study of plants natural populations growing both in isolates and in the central area of their main range is one of the most important tasks of population biology. Evolution and genetic processes occurring in plants natural populations are being effectively studied at present by molecular-genetic methods.

Birch forests along with pine forests occupy a significant part of the forested areas of Eurasia. Birches belong to the genus *Betula* and are considered to be one

of the most taxonomically problematic systematic groups. According to the recently published monograph for these species g. *Betula* includes around 65 species and subspecies [20]. Another estimation indicates 120 species of trees and shrubs in the g. *Betula* [21]. The reason for above different figures is the morphological variability of birches and the existence of a considerable number of minor birch taxa with unclear taxonomic status, which are treated by some authors as separate species and by others as varieties or forms (for example *B. carelica*, *B. carpatica*, *B. obscura*, *B. procurva*, *B. turkestanica*, *B. tianshanica* and others) The formation of many small closely related taxa is a consequence of interspecific hybridization and introgression common among birches. The systematics of birches and relationships between species can be partly resolved with the use of various molecular markers. In recent decades, due to advances in paleobotanical and molecular-genetic methods, issues of polymorphism, phylogeography, and phylogeny of birches (*B. pubescens*, *B. pendula*, *B. nana*, and *B. humilis*) have been widely discussed [1-12]. Nuclear ribosomal DNA sequences [6, 11, 13] and chloroplast DNA sequences [4, 5, 8] are mostly used to determine the phylogeny of forest- tree species. The results of these molecular genetic studies have provided new information on the evolution of the most taxonomically complex g. *Betula* as a whole, but interspecific relationships are still largely uncertain, and some data are contradictory [4, 6].

At the same time, it should be noted that the territory of Russia is still not covered by comprehensive studies of the genetic structure, genetic diversity, and the degree of interpopulation differentiation of birch populations. To

* Corresponding author: botgarden.olga@gmail.com

date, there is no clear identification of some birch taxonomic groups in the Russian Federation, such as *B. procurva*, *B. tianschanica*, and *B. turkestanica*, supported not only by morphological but also by molecular methods. The phylogenetic position and origin of these species remains unclear. The range of *B. procurva*, *B. tianschanica*, and *B. turkestanica* species is confined to the mountain system in southeast Central Asia. In addition, a group of *B. procurva* populations isolated from the main range was found to grow in the mountainous part of the Urals, in the boreal forest zone, and in the Trans-Ural forest steppe.

The aim of this work was the study of the genetic diversity and comparative analysis of the genetic structure of various birch species populations growing on the territory of the Russian Federation based on the study of cpDNA nucleotide sequences (trnL- trnF region). As well as obtaining preliminary data on the phylogenetic position and origin of *B. procurva* based on the analysis of the above nucleotide sequences. In addition, we attempted to evaluate the potential application of the trnL- trnF cpDNA region as a barcode to determine the taxonomic status and relationships between different birch species.

2 Materials and methods

Leaf material from birches young vegetative sprouts was used for this genetic study. The names of populations, geographical positions and the number of individuals sampled per population are given in Table 1.

Samples of various representatives of the g. *Betula* were collected both in natural habitats and in the cultivation areas on the territory of the Botanical Garden (*B. procurva* – Tsytsin Main Moscow Botanical Garden of Academy of Sciences). Several samples were provided in a single copy from the herbarium materials of the Botanical Garden of the Ural Branch of the Russian Academy of Sciences. *B. procurva* was represented by samples collected in Kyrgyzstan (near the rivers Osh, Gulcha), as well as growing in the National Botanical Garden. To clarify the birch species, herbarium samples collected in the Kurgan region were also used (*B. procurva*, *B. pendula*). In addition, the study included samples of birches growing in the National Botanical Garden of Kyrgyzstan named after E. Gareeva (*B. tianschanica* and *B. turkestanica*). To form a comparison group, the cpDNA sequences, publicly available in GenBank were used.

Comparative analysis of chloroplast DNA structure and differentiation parameters was performed on material obtained from both pre-dried and fresh birch leaves. Total DNA was isolated using the CTAB method [14]. For the primary analysis of the genetic variability, the intergenic spacers of cpDNA (trnL-trnF region) were used. Amplification of the regions studied was performed using universal primers, reaction conditions, and temperature regimes recommended for these regions [15,16]. Subsequent sequencing (Applied Biosystem-3130) was performed using the Sanger method with BrilliantDye v. 3.1 kit (Gene Analyzers 3130; Applied Biosystems, Foster City, CA). Editing and alignment of the obtained

sequences were performed manually in BioEdit, Mega6 [17].

Table 1. Samples geographical characteristics

No.	Species	Population/GenBank ref. No.	Number of samples
1	<i>B. procurva</i>	Main Moscow Botanical Garden of Academy of Sciences, RF, Moscow	3
2	<i>B. procurva</i>	Botanical Garden named after E.Z. Gareeva of the National Academy of Sciences, Kirgizia, Bishkek	3
3	<i>B. procurva</i>	Kirgizia, Gulcha 40°18'49"/73°26'39"	1
4	<i>B. procurva</i>	Kirgizia, Osh 40°31'41"/72°47'54"	1
5	<i>B. procurva</i>	RF, Kurganskaya oblast 55°27'/65°20'	2
6	<i>B. tianschanica</i>	Kirgizia, Gulcha 40°18'49"/73°26'39"	1
7	<i>B. turkestanica</i>	Kirgizia, Gulcha 40°18'49"/73°26'39"	1
8	<i>B. pubescens</i>	GenBank ref.№ NC 039996.1	1
9	<i>B. nana</i>	GenBank ref.№ KX703002.1	1
10	<i>B. pendula</i>	GenBank ref.№ LT855378.1	1

Since this work involves many single specimens, we grouped the specimens by their species or their geographic origin for statistical processing of the data (Table 2). Using the DNASP software package (<http://www.ub.edu/dnasp/downloadTv6.html>), we determined the level of differentiation and the distribution of genetic variability between haplotype groups (analysis of molecular variation, AMOVA). The N_{ST} , G_{ST} , and F_{ST} values were obtained to measure the molecular divergence between the species and individual groups. Molecular diversity parameters including the number of haplotypes in each group, haplotype diversity (h), and nucleotide diversity (π) were calculated using the same program. The phylogenetic relationships between the identified haplotypes of the cpDNA were reconstructed using NETWORK ver. 4.6.1.2 (Reduced Median, RM method) (<https://www.fluxus-engineering.com/sharenet.html>). The *A. incana* genotype was taken from GenBank (ref. no. NC_036752.1) for trnL-F cpDNA analysis and we used it as a combined external group.

3 Results

Sequencing of the cpDNA trnL-trnF intergenic spacer of 12 samples of various birch species yielded sequences with an average length of about 500 bp. After alignment with the sequences of the trnL-trnF intergenic spacer of *B.*

pubescens, *B. pendula*, *B. nana* downloaded from GenBank, nine haplotypes were identified, including the outer group of *A. incana*, the relationships between which were visualized as a network using the NETWORK software (Fig. 1, Table 2). The total number of sites analyzed was 526 (including gaps), the number of variable sites was 191 (including gaps), the average haplotype diversity of the total group was low (0.742), and the total number of mutations was 88.

Table 2. Group affiliation of samples, place of growth, group polymorphism parameters

Population groups	Haplo type	Hd	π	F _{ST} average for the group
<i>B.procurva</i> (Kirgizia group)	H6, H8	0,6	0,0034	0,445
<i>B.procurva</i> (Main Botanical Garden)	H1, H3	0,67	0,0019	0,285
<i>B.tianshanica</i> <i>B.turkestanica</i>	H4, H5	1	0,015	0,112
<i>B.procurva</i> (Kurgan group)	H2, H7	0	0	0,691
<i>B. pubescens</i> <i>B.nana</i> <i>B. pendula</i>	H1	0	0	0,375

During the search for unique molecular features typical for the studied species, mutations in the trnL-trnF intergenic spacer that marked certain groups of populations (synapomorphy) were found: A→C substitution in position 430, typical only for the *B. procurva* population growing in the south of Kurgan region, and the ACAAATAA (48-56) insertion typical for *B. procurva* populations from Kyrgyzstan. These mutations clearly differentiate individual groups of populations from each other, indicating potential informativeness and suitability of the selected molecular marker for study of *B. procurva* population structure.

B. procurva from the Main Botanical Garden is characterized by two haplotypes (H1 and H3, Fig. 1). The haplotype diversity value (Hd) in this group was 0.667, with a significantly low nucleotide diversity (π) (0.002). The group that combined all samples of *B. procurva* obtained from Kyrgyzstan included two unique haplotypes (H6 and H8, Fig. 1), the haplotype diversity (Hd) here was 0.600, while the nucleotide diversity (π) was minimal (0.003).

The remaining two groups: *B. procurva*, which grows in the Kurgan region and the mixed group (*B. pubescens*, *B.nana*, and *B. pendula*), are characterized by even lower values of genetic polymorphism coefficients. The number of unique haplotypes does not exceed 1, and Hd, π values are zero.

Intergroup genetic subdivision was determined by the value of fixation indices. Thus, when calculating G_{ST}, only haplotype frequencies are considered, and when calculating N_{ST} parameters, genetic divergence between haplotypes is also taken into account. The N_{ST} index (0.632±0.15) is almost two times higher than the G_{ST}

(0.390±0.38), indicating that the haplotypes of the groups studied are closely related.

The average values of the genetic diversity indices between the groups indicate active hybridization processes. A high degree of population divergence (F_{ST}) is observed between *B. procurva* groups growing in the south of Kurgan region and in Kyrgyzstan (0.727), which may be explained by their considerable geographical remoteness and hence by their long isolation. The presence of *B. procurva* growing in the south of Kurgan region in the sympatric zone cannot be excluded, which is confirmed by the significant divergence between this group and the *B. procurva* group growing in the Main Botanical Garden (0.750). Low F_{ST} value was noted between the groups of *B. procurva* from Kyrgyzstan and Main Botanical Garden (0.391) and *B. procurva* from Kyrgyzstan and *B. turkestanica*, *B. tianshanica*, having similar ranges (0.162). Zero distances were observed between groups of *B. procurva* from Kyrgyzstan and *B.pendula*, *B.nana*, *B. pubescens*, between *B.pendula*, *B.nana*, *B. pubescens* and *B.turkestanica*, *B.tianshanica* and finally between *B. procurva* from Main Botanical Garden and *B. turkestanica*, *B. tianshanica*. *B. procurva* growing in the south of Kurgan region showed low divergence values with *B.turkestanica*, *B.tianshanica* (0.286) (see Table 3).

Table 3. Genetic differentiation among population groups

Population groups		Fst	Gst	Nst
<i>B.procurva</i> (Kirgizia group)	<i>B.procurva</i> (Main Botanical Garden)	0,391	0,076	0,392
<i>B.procurva</i> (Kirgizia group)	<i>B.tianshanica</i> <i>B.turkestanica</i>	0,162	0,031	0,164
<i>B.procurva</i> (Kirgizia group)	<i>B.procurva</i> (Kurgan group)	0,727	0,411	0,727
<i>B.procurva</i> (Main Botanical Garden)	<i>B.procurva</i> (Kurgan group)	0,750	0,447	0,750
<i>B.procurva</i> (Main Botanical Garden)	<i>B.tianshanica</i> <i>B.turkestanica</i>	0,000	- 0,101	0,001
<i>B.procurva</i> (Kurgan group)	<i>B.tianshanica</i> <i>B.turkestanica</i>	0,286	0,333	0,286
<i>B. pubescens</i> <i>B.nana</i> <i>B. pendula</i>	<i>B.tianshanica</i> <i>B.turkestanica</i>	0,000	0,008	0,000
<i>B. pubescens</i> <i>B.nana</i> <i>B. pendula</i>	<i>B.procurva</i> (Main Botanical Garden)	0,000	0,000	0,000
<i>B. pubescens</i> <i>B.nana</i> <i>B. pendula</i>	<i>B.procurva</i> (Kirgizia group)	0,500	0,277	0,500
<i>B. pubescens</i> <i>B.nana</i> <i>B. pendula</i>	<i>B.procurva</i> (Kurgan group)	1,000	1,000	1,000

The spatial orientation of the cpDNA haplotypes was constructed, taking into account deletions as the fifth nucleotide type and was characterized by a star-shaped topology with a central haplotype (H1) uniting *B. pendula*, *B. nana*, and *B. pubescens*. The same haplotype (H1) was typified for the specimen *B. procurva* (Main Botanical Garden) (Fig. 1). The generalization of *B. procurva* haplotypes with *B. pendula*, *B. nana*, *B. pubescens* indicates possible existence of parental relationships between *B. procurva*, *B. nana*, *B. pendula* and *B. pubescens*, explained by the common phenomenon of chloroplast haplotype association in closely related plant species [5, 18, 19]. On the one hand, this may be the result of chloroplast capture during hybridization and, on the other hand, the presence of a common ancestor. In the latter case, the haplotypes occupy an internal position in the haplotype network, being close to the root. Despite our results, to confidently choose between these two hypotheses of common haplotypes presence in *Betula* species, it is necessary to increase the number of population samples.

The "rays" (evolutionary lines) of *B. procurva* growing in different geographical locations: Kurgan Region, Main Botanical Garden, and Kyrgyzstan departed from the central H1 haplotype. It can be assumed that *B. procurva* cpDNA is characterized by significant polymorphism, which is determined by the size of the main range and the duration of divergent groups' isolation.

Also, the obtained haplotype network confirms the origin of *B. procurva* in Central Asia with subsequent dispersal of this species to the North (mountainous part of the Urals, Trans-Ural forest-steppe). Various haplotypes of *B. procurva* cpDNA are consistent with the hypothesis of a disjunctive long-term isolated range of this species.

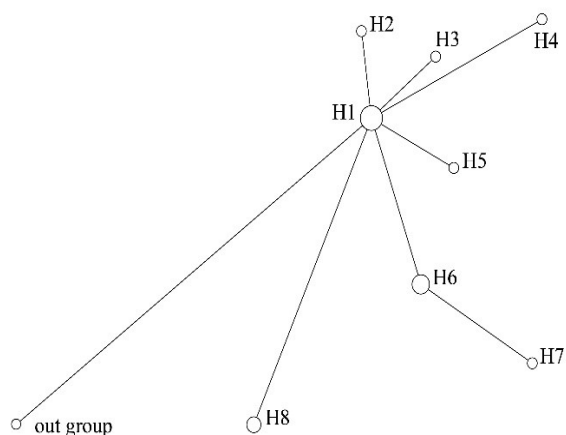


Fig. 1. Genetic network illustrating the relationships between the studied groups of the g. *Betula* constructed based on the matrix of unbiased Nei genetic distances for sequenced cpDNA sequences (trnL-trnF region)

Note: the length of the network segments corresponds to the number of mutation steps.

4 Conclusions

The primary molecular genetic study based on the comparison of chloroplast markers (trnL-F) allowed us to obtain preliminary data and put forward working hypotheses about the position of *B. procurva* in the *Betula* group. The studied genetic variability and population structure of *B. procurva*, *B. nana*, *B. pubescens*, *B. turkestanica*, *B. tianshanica* and *B. pendula* species indicate active introgression and hybridization processes. Polymorphism in all groups is significantly reduced, increasing slightly when comparing the most distant groups. The commonness of *B. procurva* cpDNA haplotypes (Main Botanical Garden) with *B. pendula*, *B. nana*, and *B. pubescens* shows that this trnL-trnF cpDNA site cannot be used for molecular identification of birch species by barcoding as a single marker, but this marker use is possible for determination of certain *B. procurva* population. Based on the cpDNA studied region (trnL-F) we revealed a clear geographic subdivision in the populations of *B. procurva* of the Trans-Urals and Central Asia.

References

1. B. Huntley, H.J.B. Birks An Atlas of Past and Present Pollen Maps for Europe: 0-13000 (1983)
2. J. Bousquet, S.H. Strauss, P. Li, Molecular Biology and Evolution **9**, 1076-1088 (1992)
3. A.K. Skvortsov, Bulletin of Moscow Society of Naturalist **107**, 73-76 (2002)
4. P. Järvinen, A.E. Palme, L.O. Morales, M. Länneppää, M. Keinänen, T. Sopanen, M. Lascoux, American Journal of Botany E **91**, 1834-1845 (2004)
5. A.E. Palme, Q. Su, S. Palsson, M. Lascoux, Mol. Ecol. **13**, 167-178 (2004)
6. J.H. Li, S. Shoup, Z.D. Chen, Rhodora **107**, 69-86 (2005)
7. M.F. Schenk, C.N. Thienpont, W.J.M. Koopman, L.W. Gilissen, M.J.M. Smulders, Tree Genetics and Genomes **4**, 911-924 (2008)
8. K. Jadwiszczak, A. Banaszek, E. Jabłonska, O. Sozinov, Tree Genetics and Genomes **8**, 1017-1030 (2012)
9. S.N. Sannikov, N.S. Sannikova, I.V. Petrova, *Essays on the theory of forest population biology*: 271 (2012)
10. Y. Tsuda, V. Semerikov, F. Sebastiani, G.G. Vendramin and M. Lascoux, Mol. Ecol (26) 589-605 (2017)
11. N. Wang, J.K. Laura, H.A. McAllister, J. Zohren, R.J.A. Buggs, Molecular Phylogenetics and Evolution **160**: 107-126 (2020)
12. N. Wang, H.A. McAllister, P.R. Bartlett, and R.J.A. Buggs, Ann Bot-London **117**, 1023-1035 (2016)
13. T. Nagamitsu, T. Kawahara, A. Kanazashi, Plant Species Biology **21**, 19-29 (2006)

14. M.E. Devey, J.S. Bell, D.N. Smith, D.B. Neale, *Theor. Appl. Genet.* (92) 673-679 (1996)
15. P. Taberlet, L. Gielly, G. Pautou, J. Bouvet, *Plant Molecular Biology* **17**, 1105-1109 (1991)
16. V. Erdogan, S.A. Mehlenbacher, *Systematic Botany* **25**, 727-737 (2000)
17. T.A. Hall, Bioedit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl. Acid Symp.* **41**, 95-98 (1999)
18. A.N. Schmidt-Lebuhn, J.M. Vos, B. Keller, E. Conti, *Mol. Phylogenet. Evol.* **65**, 23-34 (2012)
19. P.G. Wolf, R.A. Murray, S.D. Sipes, *Mol. Ecol.* **6**, 283-291 (1997)
20. K. Ashburner, H.A. McAllister, *The Genus Betula: A Taxonomic Revision of Birches*. Kew Publishing, London (2016)
21. I.Y. Koropachinskii, *Contemporary Problems in Ecology* **6**, 350-369 (2013)