

## Genetic Studies of Selected „Black-Fruit” Hawthorns: *Crataegus nigra* WALDST. et KIT., *C. pentagyna* WALDST. et KIT. and *C. chlorosarca* MAXIM.

Viktor KERÉNYI-NAGY<sup>ab\*</sup> – Tamás DEÁK<sup>c</sup> – Géza KÓSA<sup>d</sup> – Dénes BARTHA<sup>b</sup>

<sup>a</sup> Institute of Horticulture and Technology, Szent István University, Gödöllő, Hungary

<sup>b</sup> Institute of Botany and Nature Protection, University of West-Hungary, Sopron, Hungary

<sup>c</sup> Institute of Viticulture and Oenology, Corvinus University, Budapest, Hungary

<sup>d</sup> Centre for Ecological Research, Institute of Ecology and Botany, National Botanical Garden, Vácrátót, Hungary

**Abstract** – Genetic relationships of black-fruit hawthorns of the Carpathian basin have been investigated based on intergenic cpDNA sequences; first of all of the endemic *Crataegus nigra* and related species. Considerable infraspecific variation was detected in the East Asian taxon *C. chlorosarca* and some divergence in the Eurasian *C. pentagyna*. Based on the genetic analysis of investigated and reference psbA-trnH sequences, classification of sections *Crataegus* and *Sanguineae* is highly supported. From the studied taxa, *C. pentagyna* and *C. monogyna* was ordered to Sectio *Crataegus*, while *C. nigra* and *C. chlorosarca* to Sectio *Sanguineae*. Based on our data, *C. nigra* can be considered as maternal parent of the investigated *C. × degenii* hybrids.

**molekulár taxonómia / classification / intergenic cpDNA sequences**

**Kivonat** – Három kiválasztott “fekete termésű” galagonyafaj genetikai vizsgálata: *Crataegus nigra* WALDST. et KIT., *C. pentagyna* WALDST. et KIT. és *C. chlorosarca* MAXIM. Jelen tanulmányban kárpát-medencei “fekete termésű” galagonyák, elsősorban a *Crataegus nigra* és közeli rokonainak genetikai kapcsolatait vizsgáltuk intergénikus cpDNS szekvenciák alapján. Jelentős infraspecifikus variabilitást a kelet-ázsiai *C. chlorosarca* fajnál tapasztaltunk, míg az eurázsiai *C. pentagyna* esetében korlátozott számú nukleotid polimorfizmust azonosítottunk. Az általunk vizsgált és az adatbázis eredetű referencia psbA-trnH szekvenciák alapján a Sect. *Crataegus* és Sect. *Sanguineae* osztályozások támogatást kapnak. A vizsgált taxonok, a *C. pentagyna* és a *C. monogyna* a Sect. *Crataegus*ba, míg a *C. nigra* és a *C. chlorosarca* a Sect. *Sanguineae* kládokra kerültek besorolásra. Vizsgálataink alapján a *C. nigra* a *C. × degenii* hibrid anyai szülője.

**molekulár taxonómia / osztályozás / intergénikus cpDNS szakaszok**

\* Corresponding author: kenavi1@gmail.com; H-2100, GÖDÖLLŐ, Práter K. u. 1.

## 1 INTRODUCTION

The Hungarian (Danubian) hawthorn (*Crataegus nigra* WALDST. et KIT.) is an endemic species of the Carpathian Basin, occurring in the floodplains of Danube in Hungary and across the Serbian-Croatian border (for recent changes in the area see (BARTHA – KERÉNYI-NAGY, 2010). It is a mesophytic species: grows in the floodplain-forests (*Salici-Populetum nigrae* PARABUĆSKI, *Fraxino pannonicae-Ulmetum* SOÓ, *Leucojo aestivi-Crataegetum nigrae* KEVEY, FERENC et TÓTH). Morphologically, it can be identified easily, the taxon differs strongly from other taxa: it has 7–9–11–13 leaf lobes, the leaves on the upper surface are villose and the lower surface is downy, the stipules are serrate, the fruit is black with 5 seeds. *C. nigra* has 8 forms and 1 hybrid with *C. monogyna* JACQ., this hybrid, *C. × degenii* ZSÁK is also endemic in the Carpathian Basin (BARTHA – KERÉNYI-NAGY, 2010). The Degenhawthorn and its relationship to other native Hungarian hawthorn taxa were not yet studied by genetic methods. This is the first attempt to analyse the proposed origin of *C. × degenii* (syn. *C. × lambertiana* LANGE) genetically.

The more xerophytic species *C. chlorosarca* MAXIM. in the Far East is vicariant of *C. nigra*. Both species belong to Sectio *Sanguineae* ZABEL ex C. K. SCHNEID. Series *Nigrae* (LOUDON) RUSSANOV. *C. chlorosarca* is morphologically similar to *C. nigra*. It differs however from *C. nigra* with its black-purple fruit and the sepal edges have some teeth (CINOVSKIS, 1971). Genetic relation of *C. nigra* and *C. chlorosarca* has been proved earlier by LO et al (2009). The same study has shown, that East Asian and European taxa show clear divergence based on combined nuclear and chloroplast DNA data with only few exceptions – including *C. nigra*.

Small-flowered black hawthorn (*C. pentagyna* WALDST. et KIT.) has been described from the Carpathian Basin as well, but its distribution area ranges from the southern part of Banat, onto the Balkan Peninsula and to Asia Minor. It differs from *C. nigra* in morphology (3–5 leaf lobes, the leaves are leathery, the upper surface is bright and glabrous, sometimes pubescent too; the lower surface is often lanate-tomentose only in the vein-axils, the stipule is entire, the fruit is black with 5 seeds) (BARTHA – KERÉNYI-NAGY, 2010, *Table 1.*) and belongs to Sectio *Crataegus*. *C. pentagyna* is ordered to the outer basal position of Sectio *Crataegus* (LO et al., 2009). Although *C. pentagyna* and *C. nigra* show different morphological and ecological characteristics, reviewing herbarium specimens reveals that they are confused frequently in the literature.

Based on our earlier morphological studies (BARTHA – KERÉNYI-NAGY, 2010), we would like to answer the following questions with molecular markers: (1) is there any infraspecific divergence in cpDNA sequences of black-fruit hawthorns; (2) is the genetic relatedness of *C. nigra* and *C. chlorosarca*, and the classification of these taxa into the same section supported despite the geographic distance between the distribution areas; (3) are *C. pentagyna* and *C. nigra* different on the genetic level; and (4) whether the parent species of *C. × degenii* (*C. nigra* and *C. monogyna*) were identified correctly.

## 2 MATERIALS AND METHODS

### 2.1 Plant samples

We collected the samples in Hungary from natural habitats (4 samples of *C. nigra*: Szigetújfalu, 5K, compartment 6AB; 2 samples of *C. × degenii*: Szigetújfalu, the road between the compartments 4A–5B; 2 samples of *C. monogyna*: Szigetújfalu, the border of 5K–6AB compartments) or from the live collection of the Institute of Botany of the Hungarian Academy of Sciences, Vácrátót; the origin of the sampled shrubs were Vladivostok (Russia) for *C. chlorosarca* (4 samples) and Bucarest (Romania) for *C. pentagyna* (6 samples).

It should be noted, that *C. × degenii* and *C. nigra* samples originated physically from the same population, where *C. nigra* was dominant with ~99% frequency.

Table 1. Differential morphological characters between the hawthorn taxa included in the current study

Character	<i>C. chlorosarca</i>	<i>C. nigra</i>	<i>C. × degenii</i>	<i>C. monogyna</i>	<i>C. pentagyna</i>
Branch	rambling		rigid, straight		
Shoot-hair	permanently downy		soarsely downy and balding	hairless	rare, small hairy
Bracteas	dentated		less dentated	integral	
Leaves shape	triangular	deltoid			
Number of lobes	7–11		5–7–9	3–5	
Fragmentation of leaves	poor		deep		
Upper surface of leaves	hairy		±hairless	hairless	rare hairy
Lower surface of leaves	downy		full of very small and rare hair		
Hairs between the veins	even		non		tuft
Basis of leaves	round	straight or wide round	wide round	wedge	wide wedge
Edge of leaves	trough densely dentated		4–6 dentated	integral or 2–3 dentated	4–10 dentated
Sepals	reflexed		erected or V-shaped		
Sepal edge	some teeths		integral		
Peduncle	full of hair		±hairless	hairless	rare hairy
Hypanthium	rare hairy				
Fruit-colour	purple-black	black	purple	red	black
Fruit-hair	hairless		±hairless		
„Seed” number	5		2–4	1	5
„Seed” position	free		whole or inpart linked with		
Habitat	xerothermic	floodplain		xerothermic	

## 2.2 DNA amplification and sequencing

DNA was extracted from young leaves stored at –20 °C using a modified CTAB method (MSZ EN ISO 21571, 2005) originally introduced by DOYLE and DOYLE (1987). Standard polymerase chain reactions were carried out in 20 µl final volume from ~30 ng template DNA

under following conditions: 2 minutes denaturation at 94 °C was followed by 30 cycles of 30 secs denaturation at 94 °C, 30 secs primer annealing at 56 °C and 1 minute elongation at 72 °C. The reaction was closed by 5 minutes chain elongation at 72 °C.

For amplification of trnL-trnF (5'-AAAATCGTGAGGGTTCAAGTC-3' and 5'-GATTTGAACTGGTGACACGAGG-3') and psbA-trnH (5'-GTTATGCATGAACGTAATGCTC-3' and 5'-CGCGCATGGTGGATTACAATCC-3') chloroplast intergenic regions, primers used by ALBAROUKI and PETERSON (2007) for hawthorns taxa were applied. Following 1.2% agarose gel-electrophoresis, single band PCR products were isolated using the Wizard PCR Clean-Up System (Promega) according to the manufacturer's instructions. Eluted PCR products were direct sequenced using traditional Sanger sequencing on the ABI 3100 (Applied Biosystems) platform using both forward and reverse primers such accessing two times coverage.

Manually curated sequences have been uploaded to the European Nucleotide Archive (accession numbers HG937792-HG937796 for psbA-trnH and HG937797-HG937801 for trnL-trnF).

### 2.3 Sequence and phylogenetic analysis

Reference sequences of taxa belonging to sections *Sanguineae* and *Crataegus* published by ALBAROUKI and PETERSON (2007) and LO et al. (2009) were fetched from GenBank for *C. nigra* (AJ853470.1), *C. wilsonii* SARG. (EF127141.1), *C. rusanovii* CIN. (EU500281.1), *C. sanguinea* PALL. ex BIEB. (EF127143.1), *C. chlorosarca* (EU682698.1), *C. nevadensis* K. I. CHR. (EU500289.1), *C. orientalis* PALL. (EU500290.1), *C. monogyna* JACQ. (AJ853465), *C. laevigata* (POIR.) DC (AJ853468), *C. heldreichii* BOISS. (EU500295.1) and *C. pentagyna* WALDST. et KIT. ex WILLD. (EF127131.1). Multiple alignments of reference and raw sequences were carried out using the ClustalW2 tool (LARKIN et al. 2007). The raw sequences were then manually curated based on the electrophoretograms and the alignment. Completely identical sequences were joined under one sample name. Phylogenetic analysis was carried out with 1.000 bootstrap replicates and the neighbor-joining (NJ) method (SAITOU and NEI, 1987).

## 3 RESULTS

From the two investigated chloroplast intergenic regions, trnL-trnF was less variable with a total alignment length of 453 bases, only two phylogenetically informative character and 8 further nucleotide substitutions or insertions, which were monotypic to one species. The 6 bp indel identified earlier between positions 99–104 (ALBAROUKI and PETERSON, 2007) remained monotypic for *C. azarolus* L. var. *aronia* L. New polymorphic sites were identified at position 62 of the alignment, where a 1 bp deletion was recognized exclusively in the two *C. × degenii* specimens; and at position 134, where a G/T single nucleotide polymorphism (SNP) was identified, T being monotypic for *C. pentagyna*.

The psbA-trnH intergenic region, although only 298 bases long, proved to be more polymorphic with 3 phylogenetically informative character, 2 monotypic SNP and a hypervariable region. A new T/A SNP was identified at position 259 of the alignment (alignment positions are based on positions published by ALBAROUKI and PETERSON, 2007), where A is monotypic to *C. pentagyna*. Based on the sample set investigated by ALBAROUKI and PETERSON (2007), the authors proposed four indel regions between positions 130 and 190 of the alignment. In our sample set, this region of the alignment proved to be highly variable (*Figure 1*), which makes clear interpretation challenging.

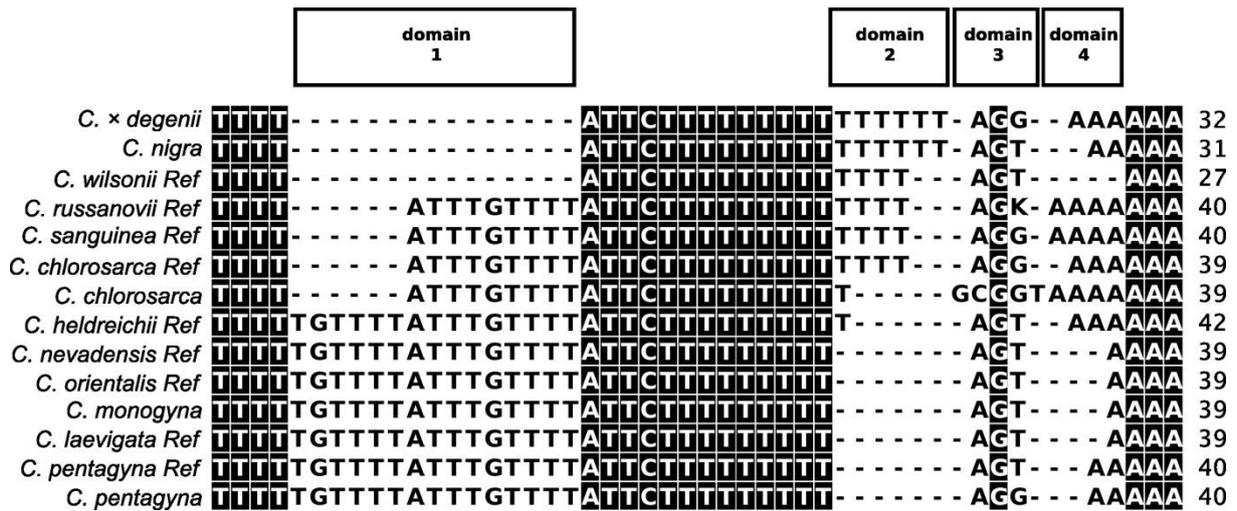


Figure 1. Alignment of the hyper-variable (HV) region of the *psbA-trnH* chloroplast intergenic region. Shaded background denotes conserved positions of the alignment. Boxes above the sequences identify polymorphic domains 1-4 discussed in the text. The HV region starts with position 134 based on the positions published by ALBAROUKI and PETERSON (2007).

First domain of the hyper-variable (HV) is monotypic in section *Crataegus*, while it is variable in *Sanguineae*, showing infraspecific variability in the case of *C. chlorosarca*. This first domain is missing from *C. nigra*, *C. × degenii* and *C. wilsonii* completely.

Second domain of the HV region is a T mononucleotide repeat, which is less informative and in this case the opportunity of sequencing errors is high. We didn't observe any infraspecific variation in this domain.

The third domain has two main characteristics. There is a GCGGT motif monotypic for all investigated *C. chlorosarca*, but not for the reference *C. chlorosarca* samples nor any other taxa. The second motif is a G/T SNP, which seems to be highly variable (data not shown). *C. rusanovii* and *C. dahurica* sequences submitted by LO et al. (2009) having an ambiguous characters at this position, the reference and the investigated Hungarian *C. pentagyna* samples have different states at this position. This is also the one and only of the investigated nucleotide position, where sequences from *C. nigra* and *C. × degenii* samples are differing. Domain four of the HV region is built up from an A mononucleotide repeat. Similar to domain two, it is less informative and error-prone. Because of possible ambiguities, domain two and four, further the G/T SNP motif of domain three were excluded from further analysis.

Intraspecific variation wasn't reported earlier for plastid intergenic sequences of hawthorns except for ambiguous bases in the GenBank entries (domain 3 on Figure 1), while the present study has shown two cases of such variation. Involved *C. pentagyna* specimens didn't differ in the investigated DNA sequences, but they showed the same T/G polymorphism compared to the reference sequences. The only remarkable infraspecific polymorphism was detected in the case of *C. chlorosarca*. In the HV region the *psbA-trnH* alignment *C. chlorosarca* specimens from Vladivostok and the reference sequences are differing on two different domains (domain 1 and 3, Figure 1).

Neighbor-joining tree of the sequenced and reference accessions show a clear distinction between the two included sections: *Crataegus* and *Sanguineae*. In section *Sanguineae* *C. nigra* and *C. × degenii* are located on the same clade, while the sequenced and reference *C. chlorosarca* samples have been ordered to different clades. The clade of section *Crataegus* has a high bootstrap support, but taxa on this clade are not structured any further.

#### 4 DISCUSSION

Based on the neighbor-joining (NJ) tree of the investigated and reference psbA-trnH sequences (Figure 2) classification of sections *Crataegus* and *Sanguineae* is highly supported.

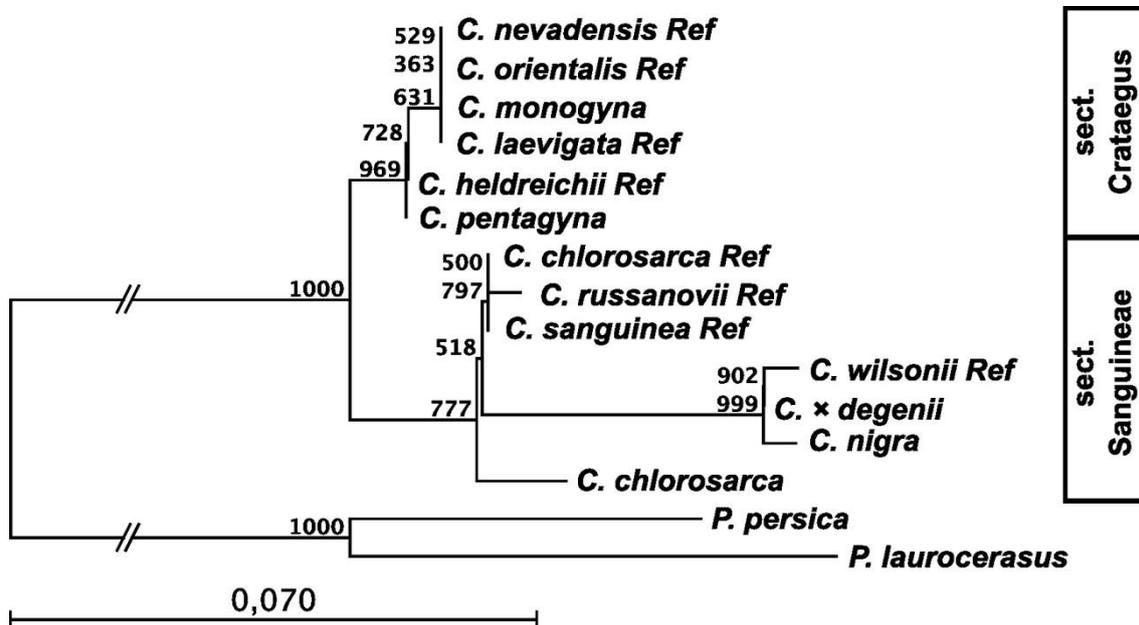


Figure 2. Neighbor joining tree of the psbA-trnH chloroplast intergenic sequences. 'Ref' denotes reference sequences fetched from the GenBank. In the case of *C. nigra*, *C. monogyna* and *C. pentagyna* no differences were found between our results and the reference sequences. *Prunus persica* and *P. laurocerasus* were used as outgroup. Numbers indicate bootstrap support from 1000 replicates.

*C. pentagyna* is correctly ordered to section *Crataegus*, in accordance to the position proposed by LO et al. (2009). The position of *C. pentagyna* is also supported by PHIPPS et al. (2003), who ordered this taxon in an own Series *Pentagynae* (C.K. SCHNEIDER) RUSSANOV. Our results confirmed the genetic differences between *C. pentagyna* and *C. nigra*. The misidentification of *C. pentagyna* in herbarium specimens can be most probably explained by the simple fact, that both hawthorns have black fruits with five seeds.

*C. chlorosarca* and *C. nigra* were ordered to the clade of Sectio *Sanguineae*. Intraspecific sequence diversity of *C. chlorosarca* can be detected also on the tree, where the investigated and reference specimens are located on different clades. The sample from Vladivostok analysed in this study shows a more distinct position compared to the *C. chlorosarca* specimen investigated by LO et al. (2009). Diversity of *C. chlorosarca* sequences could be possibly related to taxonomic differences or misidentification of the investigated specimens, such the infraspecific variability of *C. chlorosarca* psbA-trnH intergenic sequences need further confirmation. Although *C. nigra* is endemic to the Carpathian basin, it is ordered to section *Sanguineae* with species of mostly Asian origin. This is supported also by the literature (CHRISTENSEN, 1992; PHIPPS et al., 2003; LO et al., 2009). This genetic pattern – namely the similarities of an endemic from the Carpathian basin and taxa from East Asia – isn't unique. Similar phenomenon can be observed between *Syringa josikaea* JACQ. (LENDVAY et al., 2012), which is endemic in the Carpathian basin and the East Asian taxa of the *Syringa* genus. Sectio *Sanguineae* might have had earlier an Eurasian distribution, which later retreated to Asia. *C. nigra* could be considered as a relic of this former distribution.

Based on our data, *C. nigra* could possibly be the maternal parent of the investigated *C. × degenii* hybrids. The other parent *C. monogyna* could not be proved, as both hybrid samples were of same hybridization direction.

One of the main goals of this study was to clarify, whether genetic relationships of the East Asian *C. chlorosarca*, the Carpathian Basin endemic *C. nigra* and one hybrid taxon of the latter, *C. × degenii* are coherent with the high morphological similarities of these species. On the neighbor-joining tree of the investigated taxa and sequences from GenBank (Figure 2) high similarity can be observed between *C. nigra* and its hybrid, *C. × degenii* with 100% bootstrap support.

**Acknowledgements:** The research was supported by the Research Centre of Excellence „17586-4/2013/TUDPOL Szent István University”, the project 4.2.1/B-09/1/KONV–2010–0006 „Szellemi, szervezeti és K+F infrastruktúra fejlesztés a Nyugat-magyarországi Egyetemen””KTIA\_AIK\_12–1–2012–0012” and „EU – Hungary joint project, co-financed by the European Social Fund TÁMOP 4.2.4.A/2-11-1-2012-0001 ‘National Excellence Program’” and TÁMOP–4.2.1/B-09/1/KONV–2010–0006 „Szellemi, szervezeti és K+F infrastruktúra fejlesztés a Nyugat-magyarországi Egyetemen”.

## LITERATURE

- ALBAROUKI, E. – PETERSON, A. (2007): Molecular and morphological characterization of *Crataegus* L. species (*Rosaceae*) in southern Syria. — *Botanical Journal of the Linnean Society* 153(3): 255–263.
- BARTHA D. – KERÉNYI-NAGY V. (2010): Fekete galagonya – *Crataegus nigra* WALDST. et KIT. — *Tilia* 15: 54–74.
- CHRISTENSEN, K. I. (1992): Revision of *Crataegus* Sect. *Crataegus* and Nothosect. *Crataeguineae* (*Rosaceae-Maloideae*) in the Old World. — *Systematic Botany Monographs* vol. 35. The American Society of Plant Taxonomists, pp. 1–199.
- CINOVSKIS, (1971): *Crataegi Baltici* – Riga, 385 pp.
- DOYLE, J. J. – DOYLE, J. L. (1987): A rapid DNA isolation procedure for small quantities of fresh leaf tissue. — *Phytochemistry Bulletin* 19(1): 11–15.
- LARKIN, M. A. – BLACKSHIELDS, G. – BROWN, N. P. – CHENNA, R. – MCGETTIGAN, P.A. – MCWILLIAM, H. – VALENTIN, F. – WALLACE, I. M. – WILM, A. – LOPEZ, R. – THOMPSON, J. D. – GIBSON, T. J. – HIGGINS, D. G. (2007): Clustal W and Clustal X version 2.0. — *Bioinformatics* 23 (21): 2947–2948.
- LENDVAY B. – PEDRYC A. – KADERIET J. W. – WESTBERG E. – KOHUT E. – HÖHN M. (2012): A Jósika-orgona (*Syringa josikaea* JACQ. fil. ex RCHB.) aktuális és történeti biogeográfiája *Kitaibelia* 17 (1): 36.
- LO, E. Y. Y. – STEFANOVIĆ, S. – CHRISTENSEN, K. I. – DICKINSON, T. A. (2009): Evidence for genetic association between East Asian and western North American *Crataegus* L. (*Rosaceae*) and rapid divergence of the eastern North American lineages based on multiple DNA sequences. — *Molecular Phylogenetics and Evolution* 51: 157–168.
- MSZ EN ISO 21571:2005: CTAB based DNA extraction method. Hungarian Standard.
- PHIPPS, J. B. – O’KENNON, R. J. – LANCE, R. W. (2003): Hawthorns and medlars. Royal Horticultural Society Plant Collector Guide, Cambridge, pp. 1-139.
- SAITOU, N. – NEI, M. (1987): The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* 4(4): 406–425.

