STUDIA UNIVERSITATIS BABEŞ-BOLYAI BIOLOGIA, LIX, 1, 2014 (p. 55-62)

MOLECULAR EVIDENCE FOR THE HYBRID ORIGIN OF *HEPATICA TRANSSILVANICA* (RANUNCULACEAE) BASED ON NUCLEAR GENE SEQUENCES

LÁSZLÓ BARTHA[1,2,✉], KUNIGUNDA MACALIK[2] and LUJZA KERESZTES[2]

SUMMARY. Allopolyploidy (hybridisation followed by genome doubling) has been recognised as a major force driving plant speciation. The genus *Hepatica* includes both diploid and polyploid species where the origin of polyploids has not been fully established, yet. In particular, the origin of Romanian endemic, tetraploid *Hepatica transsilvanica* Fuss. remained challenging because a previous study found its incongruent placement between plastid and nuclear phylogenies (suggestive of its hybrid origin). In this study a more direct method was applied in order to shed lights on the hybrid origin of species. A fragment of nuclear *At103* gene was sequenced in *H. transsilvanica* and in both of its putative diploid progenitors, the European distributed *H. nobilis* Schreb. (var. nobilis) and the Central Asian endemic *Hepatica falconeri* Thomson. Direct *At103* sequence of *H. transsilvanica* clearly showed an additive pattern between the parental sequence types, supporting the allopolyploid origin of species. A few additional additive polymorphic sites (i.e. superimposed peaks) neither supporting, nor contradicting the hybrid origin were also found in the sequence of *H. transsilvanica* but these were not shared by all samples analysed. Origin of this 'inconclusive sequence variation' can be explained by various phenomena, like random sorting of ancestral polymorphism or paralogy. In this study a new platform is provided on which the auto- vs. allopolyploid origin of the rest of tetraploid *Hepatica* taxa can be tested.

Keywords: allopolyploidy, At103 gene, Carpathian endemic.

[1] *Institute for Interdisciplinary Research in Bio-Nano-Sciences, Babeş-Bolyai University, 42 A. Treboniu Laurean Street, 400271 Cluj-Napoca, Romania.*

[2] *Faculty of Biology and Geology, Babeş-Bolyai University, 5-7 Clinicilor Street, 400006 Cluj-Napoca, Romania.*

✉ ***Corresponding author**: **László Bartha**, Institute for Interdisciplinary Research in Bio-Nano-Sciences, Babeş-Bolyai University, 42 A. Treboniu Laurean Street, 400271 Cluj-Napoca, Romania. E-mail: lbartha.ubbcluj@yahoo.com.*

Introduction

Allopolyploidisation (hybridisation followed by genome duplication) has been considered as one of the most important processes driving plant speciation (Grant, 1981; Soltis and Soltis, 2009). Exploring patterns of hybrid speciation via sequencing requires the use of biparentally inherited nuclear markers. Identity of hybrid progenitors can be inferred if direct nuclear sequences of hybrid species show additive pattern between the parental sequence types (Campbell *et al.*, 1993). In cases when such parental sequence types are different in length and hamper direct sequencing, they can still be retrieved from hybrid species by cloning. Concerted evolution may homogenise copies of a nuclear DNA region in a hybrid species towards one of the parental (maternal or paternal) sequence type as commonly occurring in the case of internal transcribed spacer (ITS) region of the nuclear ribosomal (nr) DNA (Álvarez and Wendel, 2003). In such cases the hybrid origin can still be inferred based on conflicting topologies between phylogenies based on plastid and nuclear marker systems.

The genus *Hepatica* (*Ranunculaceae*) is a small-sized genus showing interesting biogeographic distribution characterised by intercontinental and deep inner continental disjunctions (Pfosser *et al.*, 2011). Polyploidisation played an evident role in the evolution of genus (Weiss-Schneeweiss *et al.*, 2007). Taxonomic status of several *Hepatica* species is still under dispute and the genus lacks a well-resolved backbone phylogeny. Taxa of the genus are, however, well characterised as far as their ploidy level and morphological circumscription is considered. Traditionally two morphological groups have been recognised within the genus: the one with entirely lobed leaves and another with crenate leaves. Mabuchi *et al.* (2005) proposed that crenate leaved tetraploid (2n=28) species *Hepatica transsilvanica* Fuss, *Hepatica henryi* Nakai and *Hepatica yamatutai* Steward could be autopolyploids of the diploid and crenate leaved *Hepatica falconeri* Thomson. Interestingly, *H. falconeri* grows in Central Asia, *H. transsilvanica* is confined to the Romanian Carpathians whereas the possibly conspecific *H. henryi* and *H. yamatutai* are distributed in Eastern Asia. As an alternative to the suggested autopolyploid ancestry, they could have also raised via allopolyploidisation involving *H. falconeri* (or an extinct relative to it) as one of the parental species (Weiss-Schneeweiss *et al.*, 2007). Weiss-Schneeweiss *et al.* (2007) found incongruence between plastid and nuclear phylogenies at the level of *H. transsilvanica* although these phylogenies were poorly resolved and some of their relationships were not supported. Their findings, therefore, favored the hybrid origin of *H. transsilvanica* over the autoploid origin but the ultimate inference for an allopolyploid origin of species remained largely unfulfilled.

In an effort to test the hybrid origin of the Romanian (Carpathian) endemic *H. transsilvanica* using a 'strictly direct' method, in this study a nuclear DNA marker system was employed in *H. transsilvanica* and two of its potential progenitors, the

European distributed *H. nobilis* Schreb. (var. *nobilis*) and the Central Asian endemic *H. falconeri*. Because a limited utility of the very popular phylogenetic marker nrDNA has already been found in *Hepatica* by a previous study (Weiss-Schneeweiss *et al.*, 2007), here the use of a so-called 'low-copy nuclear gene' has been endeavoured. Low-copy nuclear genes are less susceptible to concerted evolution when compared with nrDNA (Zimmer *et al.*, 1980). For the purpose of the study the nuclear At103 gene was selected due to its short length, variability and amplifiability across different plant groups with the available universal primers (Li *et al.*, 2008; Bruni *et al.*, 2010; Désamoré *et al.*, 2012).

Materials and Methods

a. Taxon and population sampling

As mentioned previously, taxon sampling included *H. transsilvanica*, *H. falconeri* and *H. nobilis*. One sample per population was analysed in the case of each species studied. Care was taken to include samples from distantly located populations of *H. transsilvanica* from the Northern and Western part, as well as from the middle of the species range (Table 1). The two samples of *H. nobilis* originated from and outside of the Carpathian Basin, respectively. The study resorted to a recently collected single Kyrgyzian accession of *H. falconeri* because additional (herbarium) samples originating from Pakistan (and collected several decades ago) turned out to be not suitable for DNA analysis.

b. Plant material, DNA isolation, PCR amplification and sequencing

Freshly collected leaf material of the studied species was desiccated in silicagel prior to DNA extraction. Total genomic DNA was extracted using the ZR Plant/Seed DNA Kit (Zymo Research). A fragment of the nuclear *At103* gene (for putative Mg-protoporphyrin monomethyl ester cyclase) spanning approximately from the middle of exon III to the end of intron IV was amplified by polymerase chain reaction (PCR) using the forward (5'-CTT CAA GCC MAA GTT CAT CTT CTA-3') and reverse (5'-TTG GCA ATC ATT GAG GTA CAT NGT MAC ATA-3') primers by Li *et al.* (2008). PCR was performed in 25 µl reaction volumes containing 12.5 µl 2× MyTaq Red Mix (Bioline), 8.5 µl dd water, 1 µl of each primers (10 µM) and 2 µl DNA-template solution of unknown concentration. Amplification of *At103* required the following PCR reaction program: initial denaturation step at 94°C for 4 min, followed by 40 cycles of denaturation at 94°C for 30 sec, annealing at 50°C for 45 sec and extension at 72°C for 30 sec. A final extension at 72°C for 5 min was inserted before final hold at 18°C. All amplifications were performed using a Gradient Palm-Cycler (Corbett Research). The success of PCR was tested by agarose gel electrophoresis. Since amplifications did not produce non-specific products (data not shown) PCR products were directly column-purified using the PCR Purification Kit of

Jena Biosciences. PCR products were both forward and reverse-sequenced by Macrogen Inc. (The Netherlands) using the amplification primers. GenBank accession numbers for all *At103* sequences generated for this study are listed in Table 1.

c. Data analysis

At103 sequences from the same accession (obtained with forward and reverse primers) were assembled using BioEdit (Hall, 1999). Sequences were aligned manually in software MEGA5 (Tamura *et al.*, 2011) from which the variable nucleotide positions were exported to a common text file.

Results and discussion

Amplification and sequencing of selected region of *At103* gene was straightforward in all accessions analysed and resulted in clear chromatograms. The length of DNA region flanked by forward and reverse primers has been 326 bp in length and showed no length variation between species. Alignment of sequences contained 10 (3 %) variable nucleotide positions (Table 2). Most (8) of these positions were located within the partial third exon which was inferred to have a length of 226 bp within the sequenced fragment of the gene (see GenBank accessions KJ842642-KJ842647).

Most of sequenced species/accessions contained unambiguous double (i.e. 'superimposed') peaks in their *At103* chromatograms (Fig. 1) (except for the species *H. falconeri* and one accession of *H. nobilis*). Nucleotide states corresponding to double peaks were codified according to *IUPAC* codes in the sequence alignment. Two types of sequence variation can be observed in the dataset regarding their *consistency*. The first type includes single nucleotide differences consistently occurring among *H. falconeri* and the two *H. nobilis* accessions at (column) positions '82' (G/A), '106' (T/C), '211' (C/T), and '289' (C/T) (Table 2). Interestingly, so called additive polymorphic sites (APS) consistently occur at the same positions among the three accessions of *H. transsilvanica* and clearly show an additive pattern between *H. falconeri* and *H. nobilis* (Table 2). This lends support to the recognition of *H. transsilvanica* as a hybrid (allotetraploid) species having *H. falconeri* and *H. nobilis* as its progenitors.

The second type of variation concerns 'randomly' occurring single nucleotide differences (with or without additivity) in the dataset at (column) positions different from those mentioned above (Table 2). Contrary to the previously presented variation, none of these nucleotide differences were shared by all samples of *H. transsilvanica* and *H. nobilis*, respectively.

Since the advent of molecular phylogenetics, this type of information (i.e. ITS profiles with additive pattern between putative parental sequence types) have been successfully used for inference of hybrid origin across different taxonomic groups. One of the earliest examples was reported by Campbell *et al.* (1993) who found additivity

of nucleotide states in *Amelanchier* × *neglecta* (Eggelst.) Eggelst. at sites where the hypothesised parental species, *A. bartramiana* (Tausch) Roemer and *A. laevis* Wieg. differed. Some of the most recent examples include the case of heterophyllous *Ranunculus penicillatus* (Dumort.) Bab. Zalewska-Gałosz *et al.* (2014) recently found that ITS sequences of this tetraploid species exhibited a clearly additive pattern between sequences of *Ranunculus fluitans* Lam. and *Ranunculus peltatus* L. illustrating its hybrid origin.

The hybrid origin of *H. transsilvanica* is neither supported nor contradicted by the second type (i.e. 'randomly occurring') sequence variation as recognised previously. Although the presence of this inconclusive variation cannot be neglected in the dataset, the 'consistent' variation is considered here as the preponderance of evidence in supporting the hybrid origin of the target species. Various phenomena can be responsible for the presence of APS in diploid *H. nobilis* and for appearance of those APS in tetraploid *H. transsilvanica* which are apparently not related to the hybrid status of species. These can result from gene flow between populations, incomplete sorting of ancestral polymorphism or paralogy (due to independent duplication events of the gene in certain populations) (Piñeiro *et al.*, 2009). Discerning the relative contribution of these phenomena to the intra-individual polymorphism of *At103* gene in selected taxa of the genus *Hepatica* was beyond the scopes of this study. This will require in the future denser population sampling, analyses of multiple samples per population, as well as sequencing multiple unlinked nuclear loci. Whatever the reason behind the above findings, the pattern of *At103* variation would be worth exploring in a phylogeographic context. Piñeiro *et al.* (2009), for example, found a clear geographic trend within variation of alleles and certain paralogs of the nuclear gene *GapC* in *Armeria pungens* (Link) Hoffmanns. & Link.

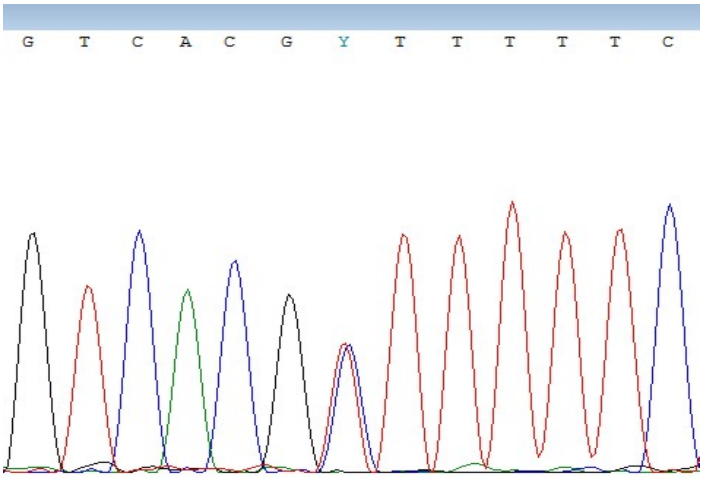

Figure 1. Unambiguous double (i.e. 'superimposed') peak in the chromatogram of *H. transsilvanica At103* sequence at nucleotide position '211' (See Table 2)

Table 1.

Information on the plant samples analysed and list of GenBank accession numbers for *At103* sequences newly generated for the study

Accession	Sampling locality	Name of collectors / suppliers	GenBank acc. numbers
Hepatica falconeri (Thomson) Steward	Kyrgyzstan, Grigorievskaya Gorge	Harry Jans	KJ842642
Hepatica transsilvanica Fuss pop. 1	Romania, Neamț County, Valea Bicăjelului	László Bartha, Bogdan-Iuliu Hurdu	KJ842643
H. transsilvanica pop. 2	Romania, Argeș County, Cheile Dâmbovicioarei Mici	László Bartha, Bogdan-Iuliu Hurdu	KJ842644
H. transsilvanica pop. 3	Romania, Hunedoara County, Hațeg	László Bartha, Kunigunda Macalik	KJ842645
Hepatica nobilis Schreb. pop. 1	Romania, Bihor County, Cetățile Ponorului	László Bartha, Attila Bartók	KJ842646
H. nobilis pop. 2	Romania, Vâlcea County, Cheile Bistriței	László Bartha, Kunigunda Macalik	KJ842647

Table 2.

Variable nucleotide positions in the *At103* sequence alignment. Nucleotide positions exhibiting additive polymorphic sites are in bold and are represented by *IUPAC* codes whereas those supporting the hybrid origin of *H. transsilvanica* are shaded in gray (Accession names as specified in Table 1)

	nucleotide positions									
accessions	73	82	106	133	145	181	184	211	253	289
H. falconeri	A	G	T	A	C	C	T	C	G	C
H. transsilvanica pop. 1	A	**R**	**Y**	**R**	C	C	C	**Y**	**S**	**Y**
H. transsilvanica pop. 2	**R**	**R**	**Y**	G	C	C	C	**Y**	G	**Y**
H. transsilvanica pop. 3	**R**	**R**	**Y**	**R**	C	**Y**	C	**Y**	G	**Y**
Hepatica nobilis pop. 1	A	A	C	A	C	C	C	T	G	T
Hepatica nobilis pop. 2	A	A	C	A	**Y**	C	C	T	G	T

Conclusions

In this study the putative hybrid origin of a polyploid *Hepatica* species has been addressed and confirmed for the first time using comparative analysis of sequences from a low-copy nuclear gene. The authors thus provide a platform on which the auto- vs. allopolyploid origin of the rest of polyploid *Hepatica* species can be tested. Future studies should address the identity of maternal and paternal species of *H. transsilvanica*, as well as its multiple vs. single origin, and – to the extent possible – the age of species.

Acknowledgements. The authors greatly acknowledge Harry Jans for collecting and providing the *Hepatica falconeri* sample analysed in the study. The help of Bogdan-Iuliu Hurdu and Attila Bartók in collecting samples of *Hepatica transsilvanica* and *Hepatica nobilis*, respectively, are also acknowledged. The authors also thank Amir Sultan for providing the Pakistani samples of *H. falconeri* from Herbarium RAW (Islamabad) for the pilot study. This work was supported by a grant of the Ministry of National Education, CNCS – UEFISCDI, project number PN-II-ID-PCE-2012-4-0595.

REFERENCES

Álvarez, I., Wendel, J.F. (2003) Ribosomal ITS sequences and plant phylogenetic inference. *Mol. Phylogen. Evol.*, **29**, 417-434

Bruni, I., De Mattia, F., Galimberti, A., Galasso, G., Banfi, E., Casiraghi, M., Labra, M. (2010) Identification of poisonous plants by DNA barcoding approach. *Int. J. Legal Med.*, **124**, 595-603

Campbell, C.S., Baldwin, B.G., Donoghue, M.J., Wojciechowski, M.F. (1993) Toward a phylogeny of *Amelanchier* (Rosaceae: Maloidae): evidence from sequences of the Internal Transcribed Spacers (ITS) of nuclear ribosomal DNA (nrDNA). *Am. J. Bot.*, **80**, 135-136

Désamoré, A., Laenen, B., González-Mancebo, J.M., Jaén Molina, R., Bystriakova, N., Martinez-Klimova, E., Carine, M.A., Vanderpoorten, A. (2012) Inverted patterns of genetic diversity in continental and island populations of the heather *Erica scoparia* s.l. *J. Biogeogr.*, **39**, 574-584

Grant, V. (1981) *Plant speciation*, Columbia University Press, New York

Hall, T.A. (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.*, **41**, 95-98

Li, M., Wunder, J., Bissoli, G., Scarponi, E., Gazzani, S., Barbaro, E., Saedler, H., Varotto, C. (2008) Development of COS genes as universally amplifiable markers for phylogenetic reconstructions of closely related plant species. *Cladistics*, **24**, 727-745

Mabuchi, T., Kokubun, H., Mii, M., Ando, T. (2005) Nuclear DNA content in the genus *Hepatica* (Ranunculaceae). *J. Plant Res.*, **118**, 37-41

Pfosser, M., Sun, B.-Y., Stuessy, T.F., Jang, C.-G., Guo, Y.-P., Taejin, K., Hwan, K.C., Kato, H., Sugawara, T. (2011) Phylogeny of *Hepatica* (Ranunculaceae) and origin of *Hepatica maxima* Nakai endemic to Ullung Island, Korea. *Stapfia*, **95**, 16-27

Piñeiro, R., Costa, A., Aguilar, J.F., Feliner, G.N. (2009) Overcoming paralogy and incomplete lineage sorting to detect a phylogeographic signal: a GapC study of *Armeria pungens*. *Botany*, **87**, 164-177

Soltis, P.S., Soltis, D.E. (2009) The role of hybridization in plant speciation. *Annu. Rev. Plant Biol.*, **60**, 561-588

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S. (2011) MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.*, **28**, 2731-2739

Weiss-Schneeweiss, H., Schneeweiss, G.M., Stuessy, T.F., Mabuchi, T., Park, J.-M., Jang, C.-G., Sun, B.-Y. (2007) Chromosomal stasis in diploids contrasts with genome restructuring in auto- and allopolyploid taxa of *Hepatica* (Ranunculaceae). *New Phytol.*, **174**, 669–682

Zalewska-Gałosz, J., Jopek, M., Ilnicki, T. (2014) Hybridization in *Batrachium* group: Controversial delimitation between heterophyllous *Ranunculus penicillatus* and the hybrid *Ranunculus fluitans* × *R. peltatus*. *Aquat. Bot.*, *in press*

Zimmer, E.A., Martin, S.L., Beverley, S.M., Kan, Y.W., Wilson, A.C. (1980) Rapid duplication and loss of genes coding for the alpha chains of hemoglobin. *PNAS*, **77**, 2158-2162