AN OPTIMIZED DNA ISOLATION PROTOCOL ENABLES AN INSIGHT INTO MOLECULAR GENETIC BACKGROUND OF ENDEMIC Moltkia petraea (Tratt.) Griseb. FROM BOSNIA AND HERZEGOVINA

Jasna Hanjalić¹*, Semir Dorić¹, Lejla Lasić¹, Belma Kalamujić Stroil¹, Selma Hasićić¹, Naris Pojskić¹

¹University of Sarajevo, Institute for Genetic Engineering and Biotechnology (INGEB), Sarajevo, B&H

Abstract

The Dinaric endemic plant species Moltkia petraea (Tratt.) Griseb. is often called a "living fossil" of ancient Tertiary flora, with great importance for Bosnia and Herzegovina’s biodiversity. Considering its narrow and limited distribution range, insufficient data on the molecular background of this species is given so far. Due to the presence of various secondary metabolites that interfere with the DNA, isolation of nucleic acids from plant cells is known to be challenging. Even in closely related species it is necessary to optimize DNA isolation protocol in order to obtain high quality PCR amplifiable DNA. We collected 91 samples from five populations in Herzegovina. Doyle and Doyle (1987) CTAB protocol was modified by adding vitamin C (ascorbic acid) to the cell lysis buffer to improve DNA yield and quality. trnL(UAA) intron and nrDNA (ITS1, ITS2) molecular markers were applied to demonstrate amplifiability of isolated DNA and elucidate the intra- and interpopulation genetic diversity. Our results suggest a successful PCR amplification for 81% of the analyzed samples. PCR-RFLP analysis of trnL(UAA) revealed that all individuals in five populations have the same haplotype based on the obtained enzymatic profile for three enzymes (TaqI, HinfI, HindII). Alignment and comparison of ITS sequences didn’t reveal any hypervariable portion that could be informative in elucidating the genetic diversity of M. petraea populations. Further studies with additional application of microsatellite loci, RAPD and AFLP methods are necessary in an attempt to get insights into the genetic diversity of M. petraea.

Key words: Moltkia petraea, trnL, nrDNA, DNA isolation

Introduction

The angiosperm family Boraginaceae includes approximately 1600 species distributed among some 110 genera, characterized by a scorpoid cymose inflorescence (Buys and Hilger, 2003). Genus Moltkia Lehm. comprises five species (Cohen, 2014), with Moltkia petraea (Tratt.) Griseb growing mostly in Mediterranean and Submediterranean regions (Kremer et al., 2016). This species is a dense, dwarf shrub, up to 40 cm high, with deep violet–blue tubular flowers, in simple forked or branching revolute cymes,
bloomimg from spring well into the summer (Fernandes, 1972; Polunin, 1987). It thrives on calcareous cliffs, in rock crevices, with altitude distribution ranging from near the sea level to 1500 m a.s.l., although it has been recorded at 2000 m a.s.l. on Mt. Durmitor in Montenegro (Trinajstić, 1974). While *M. petraea* withstands wide range of ecological parameters, it finds the optimal conditions at altitudes between 300 and 900 m a.s.l. (Marković, 1994). It is probably an ancient or even relict species (Turrill, 1929).

The current presence of *M. petraea* in gorges is probably due to its preservation within these protected areas during the last Ice Age (Georgiou et al., 1999). This species is considered an element of the endemic Ilyric-Balkan flora restricted to Bosnia and Herzegovina, Albania, Montenegro, Croatia, Macedonia and Greece (Jerković et al., 2017). In B&H, it is present in the mountain regions: Prenj, Čvrsnica, Plasa, Čabulja, Velež, Žaba, Bijela Gora and Orijen (Šoljan et al., 2009).

The only effort at surveying population (genetic) diversity of *M. petraea* has been made by Šamec (2013), who used RAPD markers to analyze *M. petraea* from Croatia. However, the author reported that any attempt at getting an amplicon failed. So far, *M. petraea* from Bosnia and Herzegovina has only been analyzed for the influence of soil properties on phenolic compound accumulation (Kremer et al., 2016) but no data on its genetic diversity exist.

Critical step for molecular characterization of any plant species is DNA isolation. Generally, it is quite difficult to extract and purify high-quality DNA because of high content of secondary metabolites (tannins, alkaloids, and polyphenols), polysaccharides and proteins in plants that mostly act as DNA polymerase inhibitors. These compounds precipitate along with DNA, degrade its quality, reduce yield and often render the sample non-amplifiable and DNA unstable for longer period (Sarwat et al., 2006, Hassan et al., 2012). As these compounds are ubiquitous in plant DNA extracts, numerous isolation protocols have been modified and used in various combinations. Due to the chemical heterogeneity among species, even closely related species may require different isolation protocols (Sharma et al., 2002).

The objectives of this study were (i) to obtain high quality DNA from fresh leaves of *M. petraea* with rapid, cost efficient and simple method for its extraction and purification, and (ii) to inspect the population (genetic) diversity using the trnL and ITS molecular genetic markers.

**Materials and methods**

**Sample Collection**

The total of 91 samples were collected from five regions in Herzegovina (Table 1, Figure 1) and stored at -20 °C. For the purposes of DNA analysis, sampling procedure included: (i) taking into account the distance between individuals; (ii) sampling the branches (of each
individual) with a few leaves on it respecting the fact that it is a rare plant species; (iii) packing plant material into paper bags (each specimen separately to avoid cross-contamination).

Table 1. Absolute and relative number of individuals collected by region and altitude

<table>
<thead>
<tr>
<th>Region</th>
<th>Number of individuals</th>
<th>Elevation (m a.s.l.)</th>
<th>(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mostarska Bijela</td>
<td>8</td>
<td>600</td>
<td>9</td>
</tr>
<tr>
<td>Grabovica</td>
<td>31</td>
<td>400</td>
<td>34</td>
</tr>
<tr>
<td>Most Begića i Begovića</td>
<td>20</td>
<td>200</td>
<td>22</td>
</tr>
<tr>
<td>Tunel</td>
<td>19</td>
<td>200</td>
<td>21</td>
</tr>
<tr>
<td>Drežnica</td>
<td>13</td>
<td>200</td>
<td>14</td>
</tr>
</tbody>
</table>

DNA extraction and qualitative-quantitative analysis

Frozen and fresh plant tissue (15 mg per sample on average) was grounded using mortar and pestle. Different protocols for DNA extraction were used: CTAB protocol (Doyle and Doyle, 1987); DNeasy Plant Mini Kit (Qiagen), protocol described by Jobes et al. (1995) and eventually the optimized procedure based on CTAB. The quality of genomic DNA was analyzed using agarose gel electrophoresis. Five microliters of genomic DNA were run in 1.5 % (w/v) agarose gel, in 1x SB (Sodium borate) buffer, pH8 (Brody and Kern, 2005) and visualized under UV light after staining with Midori green (Nippon Genetics Europe). Spectrophotometry was used to determine the concentration and quality of isolated DNA (Gallagher, 1994).

Molecular markers

Chloroplast trnL and nrDNA regions were used to demonstrate amplifiability of isolated DNA and afterwards observed for the analysis of genetic diversity. Amplification of trnL (UAA) intron was performed in an optimized PCR reaction using primers described by Taberlet et al. (1991) in 15 µl reactions consisting of 1 µl of template DNA, 2 mM Tris-HCl (pH 8.0), 10 mM KCl, 0.2 µM of each primer, 0.2 mM dNTPs, 2.5 mM MgCl2 and 1 unit of TaqGold DNA polymerase (Thermo Fisher Scientific). PCR amplification was carried out in 30 cycles (5 min at 95°C, 45s denaturation at 95°C, 30s annealing at 51°C, 45s extension at 72°C and 10 min final extension at 72°C). Successfully amplified trnL (UAA) intron was digested with restriction enzymes (Taq I, HpyF3 I, Hinf I, Hind III, Hind II, Rsa I, Ecor I, Ava II, Ban I and Alu I) in order to create RFLP profiles. Digestion was performed in individual reactions according to the manufacturers' instructions (BioLabs New England).

Figure 1. Geographical location of sampling sites

nrDNA regions of ITS1 and ITS2 (with accompanying 5.8S rRNA) were amplified in separate reactions using primers previously described by White et al. (1990). Amplifications were performed in 35 µl reactions consisting of 1 µl of template DNA, 2 mM Tris-HCl (pH 8.0), 10 mM KCl, 0.2 µM of each primer, 0.2 mM dNTPs, 2.5 mM MgCl2 and 1 unit of TaqGold DNA polymerase (Thermo Fisher Scientific). PCR parameters were 3 min at 95°C, 30s denaturation at 95°C, 30s annealing at 50°C, 1
min extension at 72°C and 10 min final extension at the same temperature with the total of 30 cycles. PCR products were sequenced by Macrogen Inc. Europe as a part of their regular capillary DNA sequencing services.

Sequence and genetic data analysis

Sequence identification analysis for nrDNA from *M. petraea* was performed using the FASTA program (Pearson, 1994). BLAST network service (Benson et al., 2003) in GenBank at NCBI was used for final sequence identification, searching for the best identity and similarity scores in local databases. Sequencing reads were assembled using DNASTAR's Lasergene software EditSeq (Burland, 2000). Electropherograms were examined manually for sequencing errors. Multiple sequence alignment analyses for nrDNA sequences were performed using ClustalW Ver.1.6 (Thompson et al., 2011) under default parameters. MSA analyzed sequences and outputs were optimized using Jalview 2.9.0b2 (Waterhouse et al., 2009) and edited by Bioedit v5.09 (Hall, 1999).

Analysis of genetic data included the generation of haplotypes based on PCR-RFLP (site polymorphism). ITS sequence variation within and among observed population, intra- and interpopulation nucleotide diversity, and nucleotide differentiation were estimated using MEGA6 (Nei, 1987; Nei and Kumar, 2000).

Results and discussion

Isolation of nucleic acids from plant cells is known to be challenging due to the presence of various inhibiting substances that bind and/or co-precipitate with them, interfere with the DNA isolation procedure and downstream applications such as PCR-based methods (Salzman et al., 1999). Thus, it is very difficult to create a universal isolation procedure and in most cases optimization is necessary. The content of biologically active compounds in plants depends on various factors: influence of soil traits, plant age, harvesting time, tissue type etc. (Končić et al., 2010, Kremer et al., 2016, Kolodziej and Sugier, 2013). Kremer et al. (2016) analyzed biochemical properties of *M. petraea* and reported the highest content of polyphenols, flavonoids and phenolic acids in leaves, the highest content of tannins highest in flowers, with the lowest content of the analyzed bioactive compounds in stems.

Šamec (2013) aimed to perform phytochemical and genetic analyses in four endemic species, i.e. *Teucrium arduini, Moltkia petraea, Micromeria croatica* and *Rhamnus intermedia*. Even after extensive optimization of DNA isolation protocols and application of different approaches, the author was not able to obtain an amplicon for any of the selected markers. UV and CD specter of DNA samples from *M. petraea* revealed the presence of unidentified compound with the maximum absorbance at 260 nm, overlapping the DNA specter.

In this study we applied several isolation protocols to obtain high quality DNA from leaves of *M. petraea* (Figure 2). Extraction using CTAB method reported by Doyle and Doyle (1987) gave low yield of DNA (Figure 2A), but samples were non-amplifiable. Application of the DNeasy Plant Mini Kit (Qiagen) (Figure 2B) and isolation protocol described by Jobes et al. (1995) (Figure 2C) resulted in DNA of even poorer quality. In order to improve DNA yield and quality, Doyle and Doyle (1987) CTAB protocol was modified. Optimization included adding vitamin C (ascorbic acid) to the cell lysis buffer (0.2%). Vitamin C is known to prevent selective binding of proteins and to have antioxidant properties. This method solved the problem of low DNA yield and co-precipitation of secondary metabolites (Figure 2D) which was proven by a successful PCR amplification for 81% of analyzed samples.
Figure 2. Electrophoretic pattern of DNA isolated by (A) CTAB protocol (Doyle and Doyle, 1987), (B) DNeasy Plant Mini Kit (Qiagen) and (C) isolation protocol described by Jobes et al. (1995). (D) Optimization of CTAB method resulted in DNA that was suitable for PCR-based analysis. The electrophoresis was performed in 1.5% (w/v) agarose gel.

**PCR-RFLP analysis**

Chloroplast trnL (UAA) intron was successfully amplified for 74 out of 91 collected samples. Digestion was performed with ten different enzymes but only three of them (TaqI, HinfI, HindIII) found recognition site within the amplicon. RFLP profiles obtained after digestion of trnL region with restriction enzymes are shown in Table 2 and Figure 3. Analysis of the restriction fragments revealed that all individuals in five observed populations have the same RFLP profile for every applied enzyme. No polymorphism (intra- or interpopulation) was detected. Consequently, haplotype distance and intrapopulation haplotype diversity equals 0.

Table 2. Restriction fragments sizes (bp) generated by digestion of the trnL region

<table>
<thead>
<tr>
<th></th>
<th>TaqI</th>
<th>HinfI</th>
<th>HindII</th>
</tr>
</thead>
<tbody>
<tr>
<td>225</td>
<td>181</td>
<td>325</td>
<td></td>
</tr>
<tr>
<td>165</td>
<td>130</td>
<td>203</td>
<td></td>
</tr>
<tr>
<td>130</td>
<td>105</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

The trnL (UAA) intron is not the most variable non-coding chloroplast region due to its catalytic function and secondary structure formations, but studies based on RFLP polymorphisms of this region have been reported in different plants (Taberlet et al., 2007; Ridgway et al., 2003; Spaniolas et al., 2010). The advantage of this molecular marker is easy cross-amplification in a large number of plants due to highly conserved primers, i.e. flanking regions. However, Taberlet et al. (2007) reported its relatively low resolution (67.3%) at the species level in plants, and it indeed proved to be insufficiently sensitive for the detection of genetic diversity within species.

Figure 3. Restriction fragments generated using restriction enzymes A) TaqI, B) HindIII, C) HinfI. The sizes of obtained fragments (given in Table 2) were determined according to ΦX174 DNA/BsuRI Ladder.

**Analysis of ITS sequences**

The sequence nucleotide composition was successfully determined for 14 of 15 sequenced samples. Length of the nrDNA region (642 bp) concurs with the reference sequence in GenBank (accession number: FJ763194). One concatenated nrDNA consensus sequence [ITSI
(260 bp) - 5.8S RNA (163 bp) - ITS2 (219 bp)] from this study was obtained and deposited in the GenBank database (accession number: KX343047).

Despite some previous evidence of ITS intrapopulation diversity in some Boraginaceae species (Kook et al., 2014), alignment and comparison of sequences (ITS1 and ITS2) obtained in this study revealed the absence of polymorphisms and substitutions, as well as lack of indel mutations.

Applied nrDNA population genetic analysis showed no intrapopulation nucleotide diversity among sequences. In relation to that, the nucleotide differentiation (Nst) equals 0, suggesting that there is no genetic differentiation between the populations based on ITS sequence polymorphisms. Notably, the absence of genetic variation within and among the populations is observed, regardless of the type of marker observed.

*M. petraea* is considered to be paleoendemic species which, according to classical Stebbins’ view (1942) are defined as depleted species, with formerly wide distribution and high genetic diversity, that had lost many or most of their biotypes, resulting in endemism (Kay et al., 2010). More often than not, restriction of habitat has negative impact on genetic diversity, causing the loss of infraspecific genetic variability. This is true for number of angiosperms (Gitzendanner and Soltis, 2000; Hamrick and Godt, 1996) such as Halacsya sendnneri (Boiss.) Dörfl., another endemic member of Boraginaceae family with distribution pattern similar to that of *M. petraea*.

Study by Coppi et al. (2014) revealed relatively low total genetic diversity (HT=0.142) in this species, based on AFLP analysis, even when compared with several case-studies on rare and/or endemic angiosperms. Although ITS region proved to be very useful in elucidating phylogenetic relationships in Boraginaceae (Cecchi and Selvi, 2009; Cohen, 2014; Chacón et al., 2016), as well as in distinguishing five species of genus *Moltki* from other Mediterranean Lithospermeae by a unique 11-bp deletion in *ITS1* (positions 235–245) (Cecchi and Selvi, 2009), it did not reveal any hypervariable portion that could be informative in elucidating the genetic diversity of populations of *M. petraea* from the locations in Bosnia and Herzegovina. Therefore, in order to get an objective insight of *M. petraea* genetic variability and to evaluate its status as a paleoendem, future studies on this species should be based on carefully selected molecular markers.

Although application of microsatellite loci as molecular markers are inevitable in plant genetic studies for diversity estimation (Vieira et al., 2016), so far no publications about detected and described microsatellite regions within the genome of *M. petraea* have been published. In such cases, it could be more suitable to apply AFLP method which is one of the most valuable tools for investigating genetic variation in plant populations because it is simple, relatively cheap, reproducible, and demands small amount of DNA which is especially convenient for the research of rare, endemic plant species. This method has been successfully applied to diploid and polyploid species from Boraginaceae family with fragmented and/or restricted distribution (Stehlik, 2003; Mengoni et al., 2006; Coppi et al., 2008).

Optionally, RAPD markers could be used, although Šamec (2013) failed to get a successful amplification in *M. petraea*, most probably because no DNA isolation protocol used in that study yielded DNA of satisfactory purity. Therefore, previous surveys on this species as well as our study showed that it is of vital importance to choose adequate approach for the
isolation of DNA since it greatly affects the success of further analytical steps.

Conclusions

Addition of vitamin C (ascorbic acid) to the cell lysis buffer solved the problem of low DNA yield, co-precipitation of secondary metabolites and provided high quality DNA useful for further molecular genetic analysis of *M. petraea*. However, no genetic diversity within and among populations of *M. petraea* in Bosnia and Herzegovina was found in this research, based on the applied markers. This may be due to the combination of these factors: (i) observed molecular genetic markers are not sensitive enough to detect intra- and interpopulation genetic variation of *M. petraea*; (ii) observed populations represent part of an old (relict) population which has highly conserved *trnL* and ITS regions. Therefore, we suggest that further investigations of this species should employ more methods for evaluation of genetic variability in endemic plants such as AFLP or RADP fingerprinting.

Acknowledgment

This research was funded by the Environmental Fund of the Federation of Bosnia and Herzegovina, grant no. 01-02-1069/2012.

References

Gitzendanner MA, Soltis PS (2000) Patterns of genetic variation in rare and widespread plant


