TÜBİTAK

Turkish Journal of Botany

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

Research Article

Turk J Bot
(2018) 42: 257-270
© TÜBİTAK
doi:10.3906/bot-1707-13

Morphological, chemical, and genetic diversity of *Gypsophila* L. (Caryophyllaceae) species and their potential use in the pharmaceutical industry

Barbara KOŁODZIEJ[1], Sylwia OKOŃ[2], Aleksandra NUCIA[2], Tomasz OCIEPA[2,*], Katarzyna LUCHOWSKA[1], Danuta SUGIER[1], Reneta GEVRENOVA[3], Max HENRY[4]

[1]Department of Industrial and Medicinal Plants, University of Life Sciences in Lublin, Lublin, Poland
[2]Institute of Plant Genetics, Breeding, and Biotechnology, University of Life Sciences in Lublin, Lublin, Poland
[3]Department of Pharmacognosy, Faculty of Pharmacy, Medical University of Sofia, Sofia, Bulgaria
[4]UMR 7565 CNRS MoBAT Team, University of Lorraine, Vandoeuvre-lès-Nancy, France

Received: 05.07.2017 • **Accepted/Published Online:** 20.12.2017 • **Final Version:** 03.05.2018

Abstract: The aim of this study was to evaluate the morphological, chemical, and genetic similarity of species belonging to the genus *Gypsophila* L. The analysis included 7 *Gypsophila* species with potential use in the pharmaceutical industry for saponin production. In order to assess the variation of the morphological traits, a number of morphological characteristics have been determined (including the number of lateral roots, diameter and total root length, average plant height, number of shoots and their diameter, number of branches on the shoot, length and width of leaves, and fresh and air-dried weight of aerial and underground parts of plants), while prosaponin content was determined as the chemical trait. Gypsogenin 3-*O*-glucuronide and quillaic acid 3-*O*-glucuronide were the dominant prosaponins determined in the *Gypsophila* roots (amounting up to 13.5 mg g^{-1} dry extract in *G. scorzonerifolia* Ser. and 12.3 mg g^{-1} dry extract in *G. acutifolia* Spreng., respectively). Moreover, a variation analysis at the molecular level was carried out based on the RAPD and ISSR methods. The conducted analysis showed that the most distinct species at the morphological, chemical, and genetic levels were *G. paniculata* L. and *G. oldhamiana* Miq.

Key words: *Gypsophila*, biodiversity, molecular markers, saponins

1. Introduction

Gypsophila L., belonging to the family Caryophyllaceae, is a genus of long-known, cultivated, ornamental, and medicinal plants (Antkowiak and Dyba, 2004). Their underground parts, containing saponins (4%–25%), were used for washing wool and silk, giving halva its fragility, and as fire extinguisher agents, while in medicine they were believed to have expectorant, laxative, and emetic properties (Henry et al., 1991; Yao et al., 2010; Korkmaz and Özçelik, 2011; Korkmaz and Dogan, 2015). Recently, a cytotoxic effect of gypsogenin 3-*O*-glucuronide derivatives towards mammalian macrophage cell lines was confirmed, and *Gypsophila* saponins are being considered as vaccine adjuvants (Gevrenova et al., 2014). The aerial parts, forming a large number of branches, ending with tiny flowers, are used for decorative purposes to adorn bouquets.

The genus *Gypsophila* includes annual, biennial, and perennial species (Barkoudah, 1962). The taxonomy is based mainly on differences in the number of filaments of stamens and pistil styles, as well as on the characteristics of sepals, petals, and capsules. The main diagnostic trait that differs among even very similar species is the presence or absence of indumentum and the size of hair on different morphological parts of the plant (Barkoudah, 1962; Falatoury et al., 2015). Currently, there are more than 150 known species belonging to the genus *Gypsophila*, making it one of the most polymorphic representatives of the subfamily Caryophylloideae and tribe Sileneae. They occur between 30° and 60° latitude in Europe and Asia, and Turkey, the Caucasus, Iraq, and Iran are considered the main centers of variation of these species (Barkoudah, 1962; Ataşlar et al., 2009). A significant portion of the species of the genus *Gypsophila* form a clade, and it seems nonmonophyletic; thus, it is a very interesting model for genetic studies (Greenberg and Donoghue, 2011). Currently, genetic similarity analyses are commonly applied in the identification of genotypes, phylogenetic evaluation, or the assessment of the relationship between many plant species (Williams et al., 1990; Hosokawa et

* Correspondence: tomasz.ociepa@up.lublin.pl

al., 2000; Li and Dao, 2011; Okoń et al., 2013). Molecular markers allow performing analysis at a very early stage of plant development, using a small amount of tissue, which in the case of many species is of great importance. Furthermore, analysis with the use of molecular markers is fast, relatively inexpensive, and not labor-intensive (Williams et al., 1990; Gupta et al., 1999; Prevost and Wilkinson, 1999).

The aim of this study was to determine the level of genetic, morphological, and chemical diversity of 7 species of the genus *Gypsophila* with potential use in the pharmaceutical industry for the production of saponins.

2. Materials and methods

2.1. Plant materials

The research material consisted of 7 species of the genus *Gypsophila* with potential use for industrial saponin production (*G. acutifolia* Stev. ex Spreng, *G. oldhamiana* Miq., *G. paniculata* L., *G. scorzonerifolia* Ser., *G. altissima* L., *G. zhegulensis* Krasnova, and *G. pacifica* Kom.). The seeds of the tested species were obtained via Index Seminum from botanical gardens (*G. acutifolia* – Botanische Gärten, Universität Bonn, Germany, IPEN-No: XX-0-BONN-3493; *G. oldhamiana* – Botanical Garden, Perm State University, Russia; *G. paniculata* – Botanical Garden UMCS, Lublin, Poland – from natural state – in Gródek n/Hrubieszów, southeastern Poland; *G. scorzonerifolia* – Botanische Gärten, Universität Mainz, Germany, IPEN-No: XX-0-MJG–19-43910; *G. altissima* – Bundesgärten Wien, Alpengarten im Belvedere, Austria, No-4010007264; *G. zhegulensis* – Botanical Garden, Perm State University, Russia; *G. pacifica* – Bundesgärten Wien, Alpengarten im Belvedere, Austria, No-4010008260).

2.2. *Gypsophila* cultivation

Gypsophila cultivation was carried out in soil of silt loam mechanical composition at the experimental farm of the University of Life Sciences in Lublin (51°09′06.81″N, 22°28′23.42″E). The experiment was designed as completely randomized blocks in three replicates of seedlings produced in multicell pallets. The plants were planted in mid-May of 2012 (*G. acutifolia*, *G. paniculata*, *G. scorzonerifolia*, *G. altissima*, *G. pacifica*) or 2013 (*G. oldhamiana*, *G. zhegulensis*), in the spacing of 0.6 × 0.5 m on plots with an area of 10 m^2.

Prior to beginning the experiment, a potassium-phosphorus fertilizer was used in the amount of 99.6 kg K ha^{-1} (in a form of potassium salt 60%) and 43.6 kg P ha^{-1} (in single superphosphate granules), and every year in early spring nitrogen fertilization was performed with 80 kg N ha^{-1} in the form of ammonium nitrate.

2.3. Morphological analysis

The analysis of the morphological features of plants was carried out at the end of the fourth or third (*G. oldhamiana*, *G. zhegulensis*) year of vegetation by digging out 20 randomly selected plants of each species and subjecting them to biometric measurements based on defining the number of lateral roots, diameter (measured at the base) and total root length, average plant height, number of shoots and their diameter (measured at the base), number of branches on the shoot, length and width of leaves, and fresh and air-dried mass of the aboveground and underground parts of plants.

2.4. Analysis of saponin

The saponin content in the underground parts of the analyzed *Gypsophila* species was determined by means of their prosaponins gypsogenin 3-*O*-glucuronide (G3G) and quillaic acid 3-*O*-glucuronide (AQ3G) using HPLC-UV as follows. One gram of powder of dried roots of seven *Gypsophila* species was mixed with 30 mL of 10% ethanol and sonicated at room temperature for 15 min (3 times). After sonication the extract was filtered by a vacuum pump (Lab Xpress filter cup, glass fiber CEM) and lyophilized. The plant extracts were dissolved in 50 mL of 0.1% trifluoroacetic acid solution (UHQ water, v v^{-1}). The purification of the saponin mixture from each extract was achieved by solid-phase extraction (SPE) on Bond Elut C18, 6 mL, 1 g cartridges (Varian, Palo Alto, CA, USA). In that case, the saponins were fixed on this support when the polar byproducts were eluted with the filtrate. The saponins were eluted with EtOH in four steps (2 mL each step) and analyzed by HPLC-UV (gradient program I).

The four EtOH elutes of the same initial extract from each species collected in the same year were poured together and evaporated to dryness (Rotavapor). The samples of saponin mixture dissolved in 5 mL of UHQ water and 2 mL of H_2SO_4 4 N were added to each test-tube. Each test-tube was heated to 100 °C for 1 h. The reaction was stopped with $NaHCO_3$ in powder to reach pH 7.2. The prosaponins were purified by SPE on Bond Elut C18, 3 mL, 500 mg cartridges (Varian). After loading samples and a washing step with water, the prosaponins were eluted from the cartridges with 5 mL of EtOH (two times). The quantification of prosaponins was performed by HPLC-UV (gradient program II).

Analytical HPLC and UPLC-ESI/MS were performed on a Shimadzu chromatographic system (Kyoto, Japan) at 210 nm and an LC-HRMS system (Thermo Fisher Scientific Inc.) equipped with a Q Exactive Plus mass spectrometer, respectively. HPLC separation was carried out on a Symmetry Shield reversed-phase column, 250 mm × 4.7 mm, 5 µm (Waters, Milford, MA, USA); UPLC-ESI/HRMS analyses were performed on a KromasilEternity XT C18, 100 mm × 2.1 mm, 1.8 µm (AkzoNobel, Malmö, Sweden). All data were acquired and processed with a Shimadzu CLASS-VP (Version 4.3) (HPLC) and Xcalibur Version 3.0 (UPLC-HRMS). The HPLC binary solvent system consisted of solvent A (sA): acetonitrile (0.1%

ortho-phosphoric acid) and solvent B (sB): 0.1% *ortho*-phosphoric acid in water. The HPLC gradient program I started with 10% sA/90% sB followed by a linear gradient to 84% sA /16% sB for 30 min. The HPLC gradient program II started with 25% sA/75% sB followed by a linear gradient to 84% sA/16% sB for 30 min. The flow rate was 2 mL min^{-1}. The binary solvent system for UPLC consisted of solvent A (sA) acetonitrile (0.1% formic acid) and solvent B (sB) 0.1% formic acid in water. The UPLC gradient program began at 25% sA and increased to 80% sA over 10 min at flow rate of 0.3 mL min^{-1}. The oven temperature was set at 40 °C for both methods. The achieved results were performed in three replications and statistically processed using variance analysis. The confidence intervals were estimated using Tukey's test at the 5% significance level.

2.4.1. Quantitative HPLC analysis and analytical performance

The G3G (96% HPLC) and AQ3G (93% HPLC) were purified from the commercial product Saponin Pure White (Merck, Germany) as described by Henry et al. (1989). The HPLC-UV analyses of both prosaponins were performed by external standard method. In the gradient program II external standard calibrations were established on five data points covering the concentration range 2.0–0.125 mg/mL for G3G and 0.5–0.125 mg/mL for AQ3G. The HPLC analyses were carried out in triplicate for each concentration and the peak area was detected at 210 nm. Calibration curves were constructed from peak areas versus analyte concentrations. The regression equations were $Y = 3355.5x - 306.16$, $r^2 = 0.9982$ (G3G) and $Y = 3886.4x - 295.24$, $r^2 = 0.9978$ (AQ3G).

The analytical performance was calculated according to the International Council on Harmonisation (ICH, 2005) recommendations (supplementary material: Table S1). Due to the lack of commercial standards, only G3G was used in the analytical performance experiments. In repeatability, reproducibility, and recovery studies, a standard solution of 0.125 mg/mL G3G was used. The repeatability and reproducibility were established by injecting the standard solution in the HPLC gradient program II 6 times over 1 day and over 10 days (3 injections per day), respectively. The sensitivity of the analytical HPLC was calculated according to the definitions of limit of detection (LOD) and limit of quantification (LOQ) using the standard deviation of the regression line of the specific calibration curve of G3G and its slope (ICH, 2005). The recovery of G3G was established by carrying out the entire SPE-HPLC procedure on a control sample (*G. paniculata*) in triplicate. The control was spiked with the standard solution of G3G. The recovery (in percentage) was calculated by subtracting the values measured for the control matrix from the sample that had been spiked with the references, divided by the added amounts of references and multiplied by 100.

The analytical HPLC-UV method was linear in the studied concentration ranges of G3G and AQ3G. The relative standard deviations (RSDs) of the repeatability and the reproducibility were established to be 1.43% and 1.97%, respectively. The LOD and LOQ were 2.7 µg/mL and 8.2 µg/mL (G3G). The recovery of G3G was 97 ± 4%. The results from the overall SPE-HPLC method were acceptable.

2.5. Morphological and chemical diversity

The Shapiro–Wilk test was applied to check the normal distribution and the means of agronomic traits of 7 *Gypsophila* species that were used for clustering analyses using a multivariate statistical package (MVSP, version 3.1) (Kovach, 1999). Cluster analysis was performed using the unweighted pair group method with arithmetic mean (UPGMA) complete linkage to group the species by the similarity of morphological and chemical parameters and plant weight.

2.6. Genetic diversity

Leaves collected from young plants and lyophilized were used for molecular analyses. DNA isolation was performed using a modified CTAB method (Doyle and Doyle, 1987).

The analysis of genetic similarity was based on two marker systems. RAPD amplification was carried out according to a modified method of Williams et al. (1990), while the ISSR method was performed according to the modified method of Ziętkiewicz et al. (1994). The reaction mixtures contained 1X PCR buffer (10 mM Tris pH 8.8, 50 mM KCl, 0.08% Nonidet P40) (Fermentas), 160 µM of each dNTP, 530 pM oligonucleotide primer, 1.5 mM $MgCl_2$, 70 ng of template DNA, and 0.5 U of Taq DNA polymerase (Fermentas) in a final reaction mixture of 15 µL. Amplification was carried out in a Biometra T1 thermal cycler programmed for 3 min at 94 °C of initial denaturation; 44 cycles of 94 °C for 45 s, 37 °C for 45 s, and 72 °C for 45 s; and a final extension at 72 °C for 7 min. A negative control was added for each run. To verify reproducibility, the primers were tested twice on the same sample.

The amplification products were separated by electrophoresis on 1.5% agarose gels containing 0.1% EtBr (1.5 h, 120 V). Fragments were visualized under a UV transilluminator and photographed using the PolyDoc System. A GeneRuler 100-bp DNA Ladder Plus was used to establish the molecular weights of the products.

The RAPD and ISSR products were scored as present (1) or absent (0) from the photographs. Only bright and reproducible products were scored. The level of polymorphism of the primer (polymorphic products/total products) and relative frequency of polymorphic products (genotypes where polymorphic products were present/total number of genotypes) (Belaj et al., 2001) were calculated.

Genetic pairwise similarities (similarity index, SI) between the genotypes were evaluated according to Dice's formula after Nei and Li (1979). A cluster analysis was conducted using the distance UPGMA method, and clustering was verified by bootstrapping. The PCA analysis was performed using PAST software. The statistical analysis was also performed with PAST software (Hammer et al., 2001).

3. Results

3.1. Morphological results

Biometric measurements indicated the variation of biometric morphological traits, characteristic of the tested species of *Gypsophila*, suggesting their different origins.

The *Gypsophila* plants formed a taproot consisting of 9 (*G. paniculata*, *G. altissima*, and *G. zhegulensis*) to 18 (*G. scorzonerifolia*) roots with an average length of approximately 35 cm and a thickness of 50.4 mm in *G. pacifica* to 88.7 cm in *G. altissima*. The average fresh and air-dried weight of single 4-year roots ranged from 434 and 117.6 g plant^{-1} in *G. oldhamiana* to 902.92 and 311.46 g plant^{-1} in *G. scorzonerifolia*, respectively. The greatest range of variation of the tested traits in the underground parts of plants was recorded in the case of air-dried mass of individual roots (coefficient of variation ranged from 10.7% for *G. oldhamiana* to 46.2% in *G. altissima*). In contrast, the differences in the distribution of other characteristics of the underground parts demonstrated low variation ($V < 20\%$) (supplementary material: Table S2).

Significant differences in the characteristics of the analyzed plants were also observed with respect to the aerial parts of *Gypsophila* plants (supplementary material: Table S3). *G. scorzonerifolia*, *G. altissima*, and *G. acutifolia* plants produced the highest and most developed aerial parts with the largest number of the thickest and highly branched shoots, while *G. paniculata* and *G. oldhamiana* formed the shortest aerial parts. *G. acutifolia* formed linear-lanceolate, long-acuminate leaves narrowed at the base and lax, dichasial, corymbose-paniculate, and glandular-hairy inflorescence, while *G. altissima* leaves were oblanceolate to linear-oblanceolate, acute to subobtuse, with inflorescence dense. The shape of *G. oldhamiana* leaves was oblong-lanceolate, they were acute and narrowed at the base, and their inflorescence was corymbose and dense, while *G. paniculata* leaves were lanceolate to linear-lanceolate, acuminate, with inflorescence paniculate-dichasial, globular with a large number of flowers. In the case of *G. pacifica* we noted oval, clasping, acute to obtuse leaves and lax inflorescence, but *G. scorzonerifolia* and *G. zhegulensis* leaves were oblong-lanceolate to narrowly ovate with apex obtuse to acute (in *G. zhegulensis* clearly narrowed at the base) and they formed stems with numerous flowers in a large, diffusely branched, leafy-bracted, panicle-like cluster inflorescence.

The average fresh and air-dried weight of shoots sampled from individual plants ranged from 352.48 and 214.82 g plant^{-1} in *G. paniculata* to 2957 and 1259.3 g plant^{-1} in *G. zhegulensis*, respectively. Similarly, with respect to the analyzed characteristics of the aerial parts of *Gypsophila* plants, the highest coefficients of variation were recorded for their air-dried mass (26.3% in *G. scorzonerifolia* to 42.5% in *G. acutifolia*) and shoot diameter (V ranged from 6.7% in *G. scorzonerifolia* and 93.7% in *G. altissima*). The lowest dispersion of the results for the aerial parts characteristics of the plants was recorded for plant height and the number of shoots per plant.

3.2. Chemical results

The roots of the analyzed *Gypsophila* species contained different amounts of saponins, the main bioactive compounds in the crude ethanol and aqueous extracts. In each extract we found many different saponins. The comparative profiling of GOTCAB saponins in the assayed species is illustrated by the extracted ion chromatograms (mass range at *m/z* 1100.00–1900.00 is consistent with GOTCAB molecular weights) (Figure 1). The GOTCAB saponins occur in complex mixtures consisting of structurally similar compounds, and their purification is a laborious and time-consuming procedure. As colorimetric methods for total saponins suffer from drawbacks (mainly the lack of specificity), two representative prosaponins of the genus *Gypsophila*, gypsogenin 3-*O*-glucuronide and quillaic acid 3-*O*-glucuronide, were determined in the studied *Gypsophila* species (Henry et al., 1989).

After a mild acid hydrolysis of the crude plant extracts, the prosaponins were purified by SPE and identified by ultraperformance liquid chromatography-electrospray ionization/high-resolution mass spectrometry (UPLC-ESI/HRMS) (Figures 2A–2E). The abundant deprotonated molecular ions $[M–H]^-$ at *m/z* 661.36 and 645.37 were first used for prosaponins recognition (Figures 2D and 2E). The fragment mass spectra MS/MS were used to identify the sapogenins and sugar residues of the prosaponins (Gevrenova et al., 2014). The glycosidic bond at C-3 of the prosaponins AQ3G and G3G showed facile cleavage in negative ion mode of ESI-MS/MS to form fragment ions at $m\ z^{-1}$ 485.33 (Figure 2D) and 469.33 (Figure 2E), respectively. These signals pointed to the loss of the glucuronic acid residue $[M–H\text{-}GlcA]^-$ and corresponded to the pseudomolecular ions of quillaic acid and gypsogenin, respectively. Furthermore, the fragmentation of $[\text{quillaic acid-H}]^-$ yielded the characteristic product ion at $m\ z^{-1}$ 405.32 by concomitant loss of $2H_2O$ and CO_2 (Figure 2D), while HCO_2H was eliminated from $[\text{gypsogenin-H}]^-$, resulting in the ion at $m\ z^{-1}$ 423.33 (Figure 2E). Thus, in our experiment, the peak G3G (Figures 2A and 2C) was identified as gypsogenin 3-*O*-glucuronide and the peak AQ3G (Figures 2A and 2B) as quillaic acid

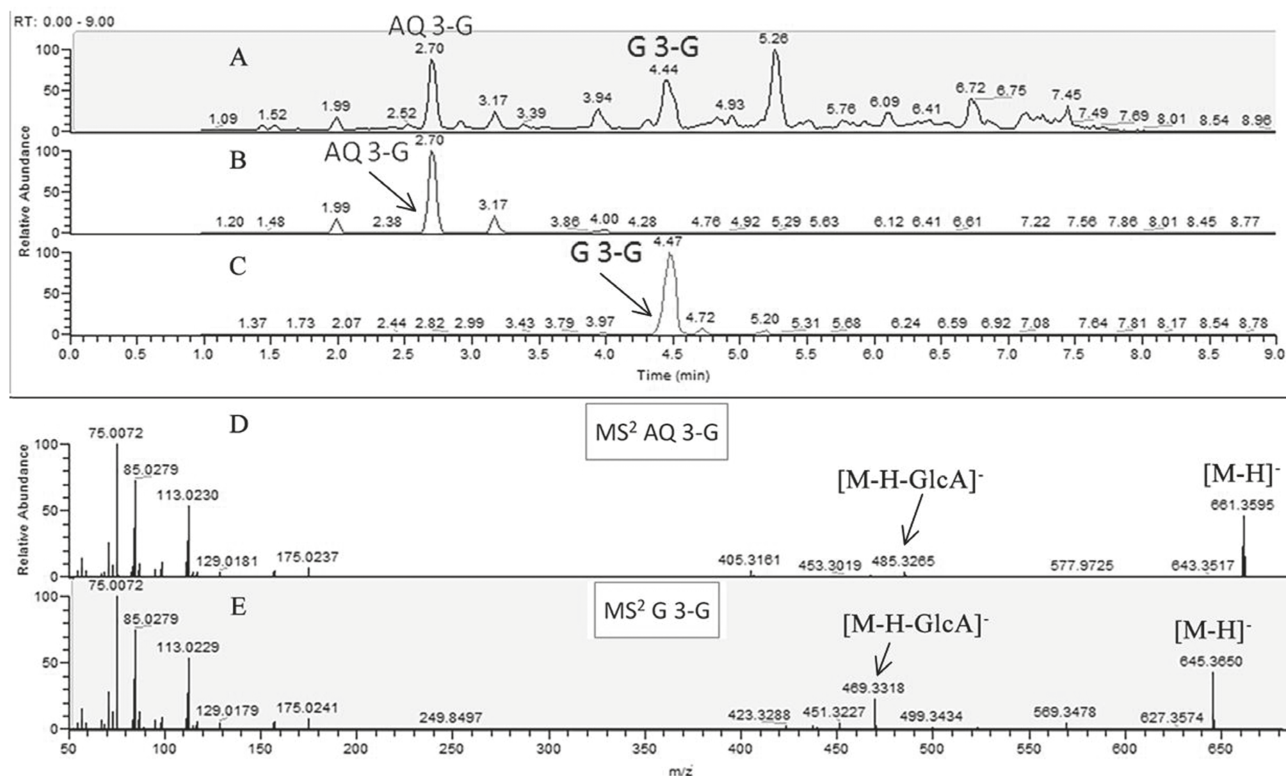

Figure 1. Extracted ion chromatograms of the GOTCAB deprotonated molecules [M-H]⁻ in the mass range at *m/z* 1100.00–1900.00 (according to the GOTCAB molecular weights). *Gypsophila scorzonerifolia* - G1, *G. acutifolia* - G2, *G. altissima* - G3, *G. pacifica* - G4, *G. paniculata* - G5, *G. oldhamiana* - G6, and *G. zhegulensis* - G7.

3-*O*-glucuronide. In this way, the presence of both prosaponins was evidenced in all *Gypsophila* extracts assayed. An RP-HPLC method involving gradient elution and UV detection for the quantification of these two prosaponins was applied following the pretreatment of the samples by SPE. The peaks of gypsogenin 3-*O*-glucuronide and quillaic acid 3-*O*-glucuronide were assessed in the HPLC chromatograms by comparing the peak retention times (Figure 2A) with those of the references, and by spiking technique as well.

In the present study, gypsogenin 3-*O*-glucuronide was the dominant prosaponin in the *Gypsophila* extracts studied (with the exception of *G. altissima*), being present in amounts up to 13.5 mg/g dry extract in *G. scorzonerifolia* (Table 1). Gypsogenin 3-*O*-glucuronide was substantially lower in both *G. altissima* and *G. paniculata* (Table 1). In contrast, the studied extracts were characterized by significantly lower amounts of quillaic acid 3-*O*-glucuronide; however, in the case of *G. altissima*, a higher amount of quillaic acid 3-*O*-glucuronide was noted. The highest content of this prosaponin was found in *G. acutifolia*, while *G. oldhamiana* and *G. zhegulensis* were the species with the lowest content of this prosaponin.

Cluster analysis of morphological and chemical traits of 7 *Gypsophila* species showed that the compared species were characterized by a rather high similarity; nevertheless, two separate clusters could be distinguished in the dendrogram. The first of them comprised *G. paniculata* and *G. oldhamiana* (approximately 92% similarity), and the remaining species were in the second one. In the second cluster, the highest percentage of similarity (approximately 95%) was observed in the case of *G. acutifolia* and *G. pacifica*, and the lowest one between *G. zhegulensis* and the remaining species (Figure 3).

3.3. Genetic analysis results

The analysis of similarities at the genetic level confirmed diverse levels of similarities between the analyzed *Gypsophila* species. As a result of the amplification reaction carried out based on the RAPD and ISSR markers, 80.31% and 95.86% polymorphic products were obtained, respectively. The frequency of polymorphic bands ranged from 43% to 100% in the RAPD method (Table 2) and from 80% to 100% in the ISSR method (Table 3). The results of the analysis of RAPD polymorphic markers and ISSR were used to create a genetic distance matrix, according to the Dice formula (SI).

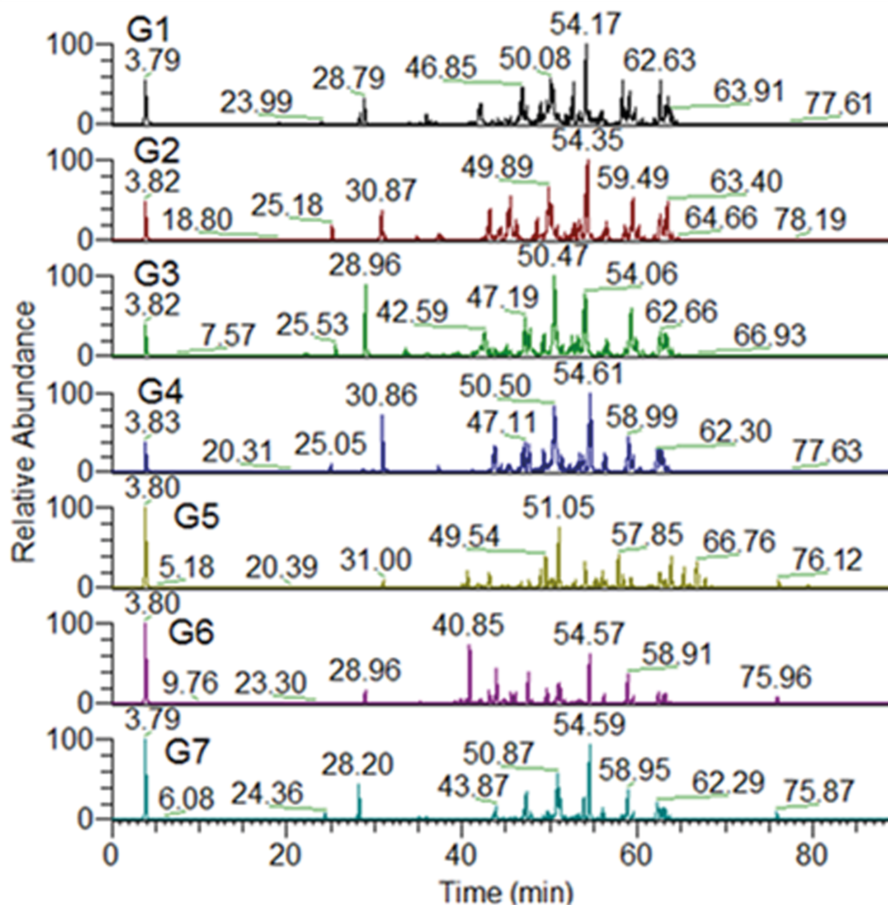

Figure 2. UPLC-ESI/HRMS chromatogram in negative ion mode of prosaponins from *G. oldhamiana* (2014) after mild acid hydrolysis. A- Total ion chromatogram (TIC); B- extracted ion chromatogram of quillaic acid 3-glucuronide (AQ3G) $[M-H]^-$ at $m\ z^{-1}$ 661.36; C- extracted ion chromatogram of gypsogenin 3-glucuronide (G3G) $[M-H]^-$ at $m\ z^{-1}$ 645.37; D- MS/MS spectrum of AQ3G; E- MS/MS spectrum of G3G.

Table 1. Saponin contents in the roots of 7 *Gypsophila* species expressed as their prosaponins: gypsogenin 3-*O*-glucuronide (G3G) and quillaic acid 3-*O*-glucuronide (AQ3G).

Species	G3G (in mg/g of root dry weight)	AQ3G (in mg/g of root dry weight)
G. scorzonerifolia	13.5 a*	5.5 b
G. pacifica	8.5 b	5.1 b
G. paniculata	4.3 c	4.1 c
G. altissima	2.8 d	4.1 c
G. acutifolia	12.0 a	12.3 a
G. zhegulensis	8.8 b	1.8 d
G. oldhamiana	6.5 c	2.3 d

*Values designated with the same letters within columns do not significantly differ at 5% error level (Tukey’s test).

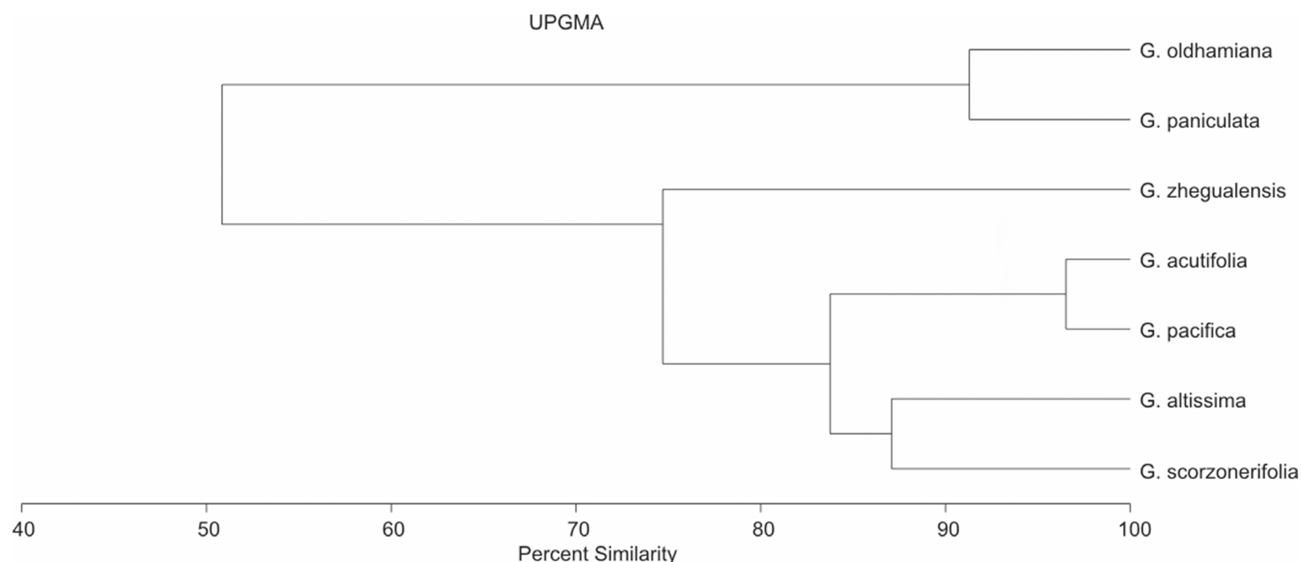

Figure 3. Cluster analysis of *Gypsophila* morphological and chemical traits in relation to the species.

Table 2. Characteristics of RAPD primers used to estimate the genetic variability of the analyzed *Gypsophila* genotypes.

Primer	Primer sequence, 5'-3'	Number of products		Primer diversity, %
		Total	polymorphic	
M16	GTAACCAGCC	8	4	50
L07	AGGCGGGAAC	12	11	92
J19	GGACACCACT	6	6	100
L11	ACGATGAGCC	9	9	100
N06	GAGACGCACA	5	4	80
G04	AGCGTGTCTG	6	6	100
L08	AGCAGGTGGA	8	7	88
N08	ACCTCAGCTC	5	4	80
K15	CTCCTGCCAA	7	3	43
L15	AAGAGAGGGG	6	5	83
L04	GACTGCACAC	3	2	67
T01	GGGCCACTCA	7	7	100
M07	CCGTGACTCA	4	3	75
C05	GATGACCGCC	5	4	80
D18	GAGAGCCAAC	9	6	67
L05	ACGCAGGCAC	9	6	67
B06	TGCTCTGCCC	6	6	100
N10	ACAACTGGGG	5	4	80
F02	GAGGATCCCT	5	4	80
N07	CAGCCCAGAG	2	1	50
Total		127	102	
Average/primer		7	6	
Average/genotype		18.14	14.57	

Table 3. Characteristics of ISSR primers used to estimate the genetic variability of the analyzed genotypes of *Gypsophila.*

Primer sequence, 5'-3'	Primer sequence, 5'-3'	Number of products		Primer diversity, %
		Total	Polymorphic	
SR01	$(AG)_8G$	10	9	90
SR11	$(AC)_8G$	20	20	100
SR16	$(CT)_8T$	12	12	100
SR22	$(CA)_8G$	13	12	92
SR28	$(TG)_8G$	11	11	100
SR37	$(AC)_8C$	11	10	91
SR69	$(AC)_8G$	13	13	100
SR31	$(AG)_8YC$	7	6	86
SR52	$(TG)_9A$	11	11	100
SR23	$(CA)_8GC$	11	10	91
SR61	$(AC)_9G$	12	12	100
SR41	$(AG)_{10}C$	16	15	94
SR58	$(ACC)_6T$	18	17	94
SR46	$(GA)_{10}A$	17	16	94
SR43	$(GT)_8A$	10	10	100
SR45	$(GA)_8T$	10	10	100
SR47	$(CA)_8A$	12	11	92
SR62	$(ATG)_6G$	15	14	93
SR63	$(ATG)_6C$	7	6	86
SR68	$(AC)_8T$	19	19	100
SR32	$(AG)_8YT$	19	19	100
SR64	$(ATG)_6AC$	11	11	100
SR75	(ATG)6 A	14	13	93
SR78	(ATG)6 G	15	14	93
SR79	(ATC)6 A	20	20	100
SR81	(ATC)6 C	13	12	92
SR82	(ATC)6 G	16	15	94
SR83	$(CA)_8RC$	17	17	100
SR84	$(CA)_8RG$	5	4	80
SR85	$(CA)_8RT$	15	14	93
SR87	$(GT)_8T$	14	14	100
SR88	$(GT)_8YA$	17	16	94
SR89	$(GT)_8YC$	14	13	93
SR90	$(GT)_8YG$	14	14	100
Total		459	440	
Average/primer		12.5	12	
Average/genotype		65.57	62.86	

As regards the RAPD method, the highest values of similarity indices was recorded between *G. scorzonerifolia* and *G. acutifolia*, which amounted to 0.951, while the lowest was found for *G. altissima* and *G. paniculata* (0.466). The average value was 0.762 (Table 4). The value of SIs based on the ISSR method ranged from 0.252 to 0.937, and the average was equal to 0.596 (Table 5). *G. acutifolia* and *G. zhegulensis* showed the highest similarity among the genotypes tested, and the largest genetic distance relative to the other genotypes was found for *G. paniculata*.

Based on the matrices of SIs, the hierarchical clustering of the tested genotypes was carried out by UPGMA method. On the dendrogram obtained based on the RAPD method, the genotypes of *G. scorzonerifolia*, *G. acutifolia*, *G. zhegulensis*, and *G. oldhamiana* were clustered together, while *G. pacifica* and *G. altissima* were located on the border of the cluster. *G. paniculata* was located on the border of the dendrogram (Figure 4).

On the dendrogram obtained based on ISSR markers, the genotypes of *G. pacifica*, *G. altissima*, *G. acutifolia*, *G. zhegulensis*, and *G. scorzonerifolia* formed a joint group of clusters. *G. oldhamiana* and *G. paniculata* were located on the border of the dendrogram (Figure 5).

Table 4. Similarity matrix for Dice coefficients of 7 species of *Gypsophila* based on RAPD markers.

	G. pacifica	*G. altissima*	*G. paniculata*	*G. scorzonerifolia*	*G. acutifolia*	*G. zhegulensis*	*G. oldhamiana*
G. pacifica							
G. altissima	0.78022						
G. paniculata	0.52632	0.46575					
G. scorzonerifolia	0.85561	0.83978	0.54305				
G. acutifolia	0.84946	0.86667	0.53333	0.95135			
G. zhegulensis	0.81081	0.86034	0.48322	0.89130	0.92896		
G. oldhamiana	0.84103	0.79365	0.55346	0.91753	0.88083	0.83333	

Table 5. Similarity matrix for Dice coefficients of 7 species of *Gypsophila* based on ISSR markers.

	G. pacifica	*G. altissima*	*G. paniculata*	*G. scorzonerifolia*	*G. acutifolia*	*G. zhegulensis*	*G. oldhamiana*
G. pacifica							
G. altissima	0.90909						
G. paniculata	0.25893	0.26027					
G. scorzonerifolia	0.85053	0.85161	0.26573				
G. acutifolia	0.86939	0.84583	0.25225	0.85350			
G. zhegulensis	0.88577	0.87526	0.25166	0.85000	0.93737		
G. oldhamiana	0.44335	0.43939	0.25556	0.46512	0.45274	0.43309	

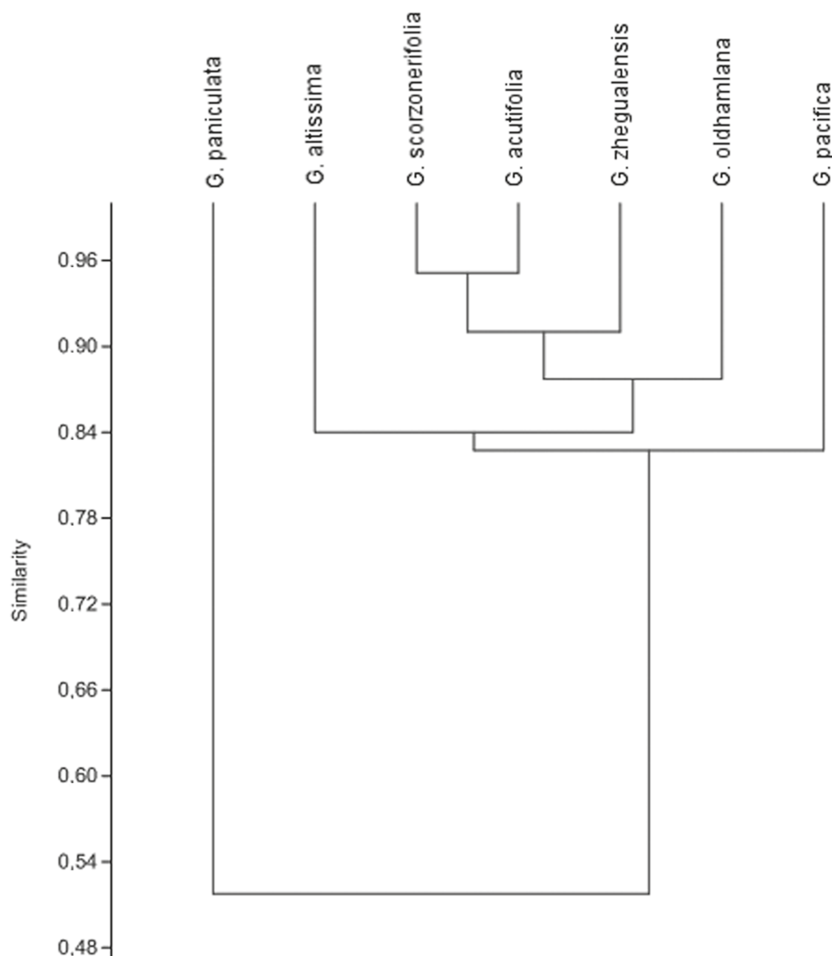

Figure 4. UPGMA dendrogram of 7 species of the genus *Gypsophila* based on polymorphism identified by the RAPD method.

4. Discussion

The assessment of biodiversity within different plant species is a very important aspect enabling the evaluation of genetic resources of a given species and its use in cultivation or plant protection. Currently, there are a number of different methods analyzing genetic variation, both within species and between different species of plants. Molecular markers have found a wide use in assessing genetic diversity; they are particularly important where assessment at the morphological level is insufficient due to a narrow gene pool and high similarity of the analyzed objects. Among the many marker systems, RAPD and ISSR, as well as the more demanding AFLP method, have become the most common research techniques (Mohan et al., 1997; Schulman, 2007). These techniques have often been used to assess the genetic variation of many plant species (Soodabeh et al., 2011; Samal et al., 2012; Alam et al., 2016; Banerjee et al., 2016; Santos et al., 2016), including protected species (Brock et al., 2007; Okoń et al., 2014; Manole et al., 2015).

The present study attempted to evaluate the diversity of 7 *Gypsophila* species at the morphological, chemical, and genetic levels.

In the literature, there is only fragmentary information about the morphological traits of *Gypsophila*, mainly related to the cultivated varieties of *G. paniculata* intended for decorative purposes. In the present study, biometric measurements were carried out of both roots and aerial parts of the analyzed *Gypsophila* species. The air-dried weight of *G. paniculata* roots was similar in our study to that reported by Antkowiak (2004), but slightly lower than in the study of Iankulov (1974). The plant heights were similar as in the studies of Iankulov (1974) and Hayashi et al. (2009), while the diameter and number of stems, as well as their fresh weights, remained at levels similar to those of the Japanese study (Hayashi et al., 2009). Cluster analysis of the morphological traits showed that *G. oldhamiana* and *G. paniculata* were characterized by the most distinct traits of the analyzed species. The differences in morphological traits of these species are related to their

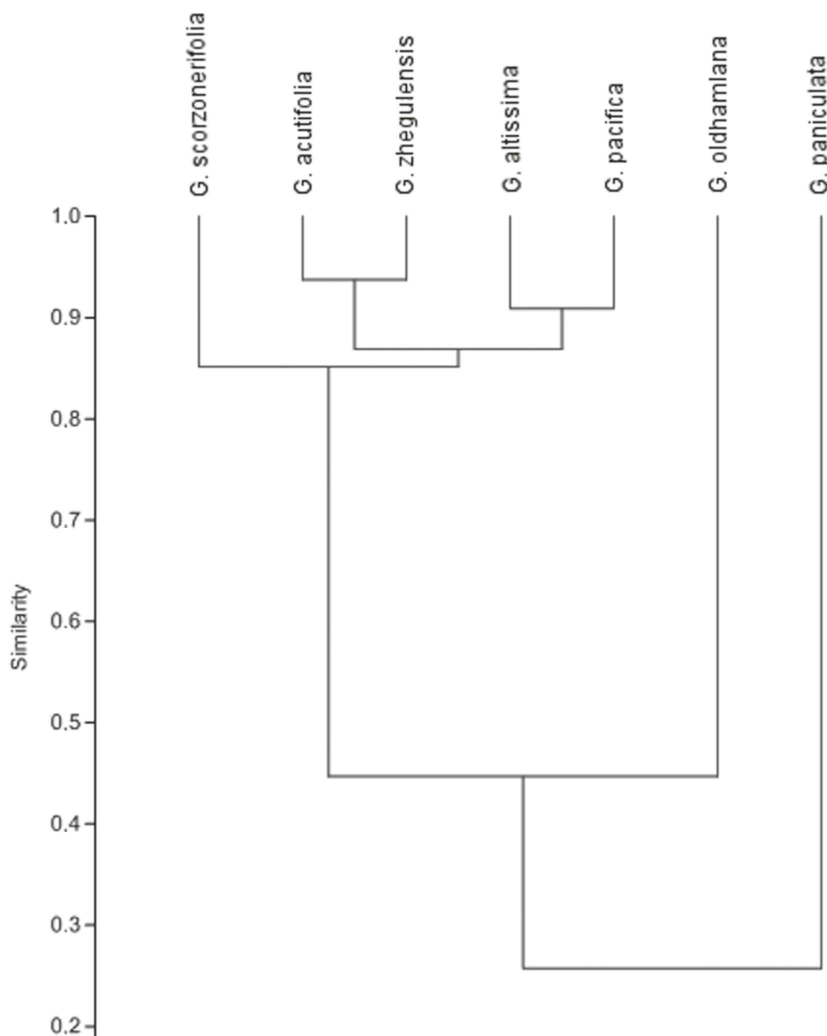

Figure 5. UPGMA dendrogram of 7 species of the genus *Gypsophila* based on polymorphism identified by the ISSR method.

origin and occurrence. *G. oldhamiana* is a species native to China, encountered in China, Mongolia, and Korea, while *G. paniculata* is a species occurring in East Europe and Asia (Barkoudah, 1962; Saxonov et al., 2011).

In line with the results of Henry et al. (1989) and Gevrenova et al. (2014), the crude extracts from 4- and 3-year-old roots of 7 *Gypsophila* species contained triterpenoid saponins from the group of glucuronide oleanane-type triterpene carboxylic acid 3,28-bidesmosides (GOTCAB). Two common saponin precursors, G3G and AQ3G, were quantified in the assayed species. Their content is important from the phylogenic point of view and refers directly to the biosynthetic saponin pathways. In this study, liquid chromatography and high-resolution Orbitrap-based mass spectrometry was used for prosaponin identification for the first time. This is also the first report on saponin/prosaponin content of the majority of the studied *Gypsophila* species, including *G. scorzonerifolia*, *G. acutifolia*, *G. altissima*, *G. pacifica*, *G. oldhamiana*, and *G. zhegulensis*.

In the past, a standard for hemolytic tests in most saponin determinations (so-called Saponin Pure White by Merck), containing about 14.3% gypsogenin 3-*O*-glucuronide (G3G), was obtained from underground parts of *G. paniculata* (Henry et al., 1989). In our investigations we also found G3G in *G. paniculata* roots, but in smaller amounts (4.3 mg g^{-1} dry weight), consistent with previous results of Henry et al. (1989), but higher than the results of Gevrenova et al. (2010, 2014). The saponin contents in *G. paniculata* roots were also similar to those found by Antkowiak and Dyba (2004) in 2- and 3-year-old plants. On the other hand, in roots of *G. scorzonerifolia* we found about 13.5 mg g^{-1} G3G and 5.5 mg g^{-1} AQ3G, many times more than in the case of in vitro cultures of that

species (Gevrenova et al., 2010). It is worth repeating that in the available literature there is no information about the saponin contents in any other *Gypsophila* species under study.

Analyses of the genetic variation of different *Gypsophila* species have already been carried out several times. Korkmaz and Dogan (2015) assessed the genetic variation of *Gypsophila* based on RAPD and ISSR markers. The focus of their research included 14 *Gypsophila* species from different phytogeographic regions of Turkey. Calistri et al. (2014) used ISSR, AFLP, TRAP, and cpSSR for the analysis of genetic distance of 5 *Gypsophila* wild species from Europe and Asia and 13 commercial hybrids with similar phenotypes, but of unknown origin. Martínez-Nieto et al. (2013) used AFLP markers to assess the genetic diversity of 2 subspecies of *Gypsophila struthium*. The present study analyzed the genetic variation of 7 *Gypsophila* species, which can be used in the pharmaceutical industry for the production of saponins. Markers used in this study allowed the achieving of a high level of polymorphic fragments, indicating a good selection of methods for the analysis of *Gypsophila* species. Polymorphic fragments obtained by the RAPD method represented 80.31% of all the amplicons; in the ISSR method, 95.86% of the fragments were polymorphic. Similarly, Korkmaz and Dogan (2015) obtained a high rate of polymorphic products for their analyzed *Gypsophila* species. Polymorphic fragments obtained by the RAPD method amounted to 92.7%, while those determined by ISSR markers constituted 93.8%. Calistri et al. (2014), using the ISSR method, also achieved a high level of polymorphic products amounting to 96.3%. The higher number of polymorphic products obtained with ISSR markers indicated the greater potential of this method; however, due to their speed and simplicity, RAPD markers are still used in the analysis of genetic variation of many species (Fernandez et al., 2002; Behera et al., 2008; Ferrão et al., 2013).

Cluster analysis based on molecular markers showed that the analyzed species exhibited different levels of genetic similarity. The frequency of polymorphic bands ranged from 0.466 to 0.951 in the RAPD method and from 0.252 to 0.937 for ISSR markers. Similar SIs were reported by Calistri et al. (2014) and Korkmaz and Dogan (2015). Analysis of the dendrograms showed that the most distinct species among the analyzed forms was *G. paniculata*. This species was also characterized by the most different morphological traits. Dendrograms derived from the polymorphisms of DNA markers also showed that the species *G. oldhamiana* differed significantly from the other species. Distinctiveness of the characteristics of this species, as in the case of *G. paniculata*, was associated with its geographic occurrence. The analyses carried out by Calistri et al. (2014) also indicated a separation of *G. oldhamiana*, which clustered on the border of dendrograms. Other species analyzed in the present study were derived from the former Soviet Union: *G. scorzonerifolia* – around the Caspian Sea; *G. pacifica* – Manchuria and the western Siberian coast between 40°N and 50°N, *G. altissima* – Ukraine, the Caucasus, Kazakhstan, the Russian Federation, Kyrgyzstan; *G. acutifolia* – Transcaucasia; *G. zhegulensis* – the Volga basin (Barkoudah, 1962; Saxonov et al., 2011). The common geographical origin of these species has been confirmed by the joint clustering in dendrograms and high similarity at both morphological and genetic levels.

The conducted analyses showed that the most different species at both morphological and genetic levels was *G. paniculata*; high morphological and genetic distinctiveness was also exhibited by *G. oldhamiana*. Therefore, it can be assumed that these species will be most appropriate in searching for genotypes containing the preferred composition of saponins. These two genotypes have very low contents of G3G and AQ3G. The most abundant in saponins were *G. scorzonerifolia*, *G. acutifolia*, *G. pacifica*, and *G. zhegulensis*, which create common cluster groups in the dendrograms based on molecular markers. These species were also similar based on their morphological traits. The high similarity of these species at the genetic level may suggest that they have some valuable traits in terms of saponin production, so we can conclude that the species derived from the former Soviet Union region could be suitable for selection of genotypes that can be used in the pharmaceutical industry.

Acknowledgment

The authors are grateful to Associate Professor Paraskev Nedyalkov, Faculty of Pharmacy, Medical University, Sofia, for his help in performing the UPLC-ESI/HRMS method.

References

Alam SMM, Siddika S, Haque ME, Islam MA, Mukherjee A, Sikdar B (2016). Genetic diversity of some upland and lowland rice cultivars in Bangladesh using RAPD, ISSR and SSR markers. Nucleus 59: 15-23.

Antkowiak W (2004). Yield of *Gypsophila paniculata* roots in dependance of different NPK fertilization. Roczniki Akademii Rolniczej w Poznaniu - CCCLX 38: 3-8 (in Polish with abstract in English).

Antkowiak W, Dyba S (2004). Effect of differentiated NPK fertilization on saponin content in roots of tall and spread gypsophyll (*Gypsophila paniculata* L. and *G. repens* L.) in the second and third years of cultivation. Prace z Zakresu Nauk Rolniczych 97: 9-14.

Ataşlar E, Potoğlu Erkara İ, Tokur S (2009). Pollen morphology of some *Gypsophila* L. (Caryophyllaceae) species and its taxonomic value. Turk J Bot 33: 335-351.

Banerjee R, Chattopadhyay S, Saha AK (2016). Genetic diversity and relationship of mulberry genotypes revealed by RAPD and ISSR markers. Journal of Crop Improvement 30: 478-492.

Barkoudah YI (1962). A revision of *Gypsophila*, *Bolanthus*, *Ankyropetalum* and *Phyrna*. Wentia 9: 1-203.

Behera TK, Singh AK, Staub JE (2008). Comparative analysis of genetic diversity in Indian bitter gourd (*Momordica charantia* L.) using RAPD and ISSR markers for developing crop improvement strategies. Sci Hortic-Amsterdam 115: 209-217.

Belaj A, Trujillo I, De La Rosa R, Rallo L, Giménez MJ (2001). Polymorphism and discrimination capacity of randomly amplified polymorphic markers in an olive germplasm bank. J Am Soc Hortic Sci 126: 64-71.

Brock J, Aboling S, Stelzer R, Esch E, Papenbrock J (2007). Genetic variation among different populations of *Aster tripolium* grown on naturally and anthropogenic salt-contaminated habitats: implications for conservation strategies. J Plant Res 120: 99-112.

Calistri E, Buiatti M, Bogani P (2014). Characterization of *Gypsophila* species and commercial hybrids with nuclear whole-genome and cytoplasmic molecular markers. Plant Biosyst 150: 11-21.

Doyle JJ, Doyle JL (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochemical Bulletin 19: 11-15.

Falatoury AN, Assadi M, Ghahremaninejad F (2015). Taxonomic significance of indumentum in the genus *Gypsophila* L. (Caryophyllaceae). Nova Biologica Reperta 2: 91-102.

Fernandez M, Figueiras A, Benito C (2002). The use of ISSR and RAPD markers for detecting DNA polymorphism, genotype identification and genetic diversity among barley cultivars with known origin. Theor Appl Genet 104: 845-851.

Ferrão LFV, Caixeta ET, de Souza F, Zambolim EM, Cruz CD, Zambolim L, Sakiyama NS (2013). Comparative study of different molecular markers for classifying and establishing genetic relationships in *Coffea canephora*. Plant Syst Evol 299: 225-238.

Gevrenova R, Joubert O, Mandova T, Zaiou M, Chapleur Y, Henry M (2014). Cytotoxic effects of four *Caryophyllaceae* species extracts on macrophage cell lines. Pharm Biol 52: 919-925.

Gevrenova R, Stancheva T, Voynikov Y, Laurain-Mattar D, Henry M (2010). Root *in vitro* cultures of six *Gypsophila* species and their saponin contents. Enzyme Microb Tech 47: 97-104.

Greenberg AK, Donoghue MJ (2011). Molecular systematics and character evolution in Caryophyllaceae. Taxon 60: 1637-1652.

Gupta PK, Varshney RK, Sharma PC, Ramesh B (1999). Molecular markers and their applications in wheat breeding. Plant Breeding 118: 369-390.

Hammer Ø, Harper DA, Ryan PD (2001). PAST: Paleontological statistics software package for education and data analysis. Palaeontol Electron 4: 5-7.

Hayashi T, Kakeya H, Konishi K (2009). Path analysis of interrelations of some components determining the number of florets per shoot in *Gypsophila paniculata* L. J Jpn Soc Hortic Sci 78: 330-334.

Henry M (2005). Saponins and phylogeny: example of the "Gypsogenin group" saponins. Phytochem Rev 4: 89-94.

Henry M, Pauthe-Dayde D, Rochd M (1989). Extraction and high-performance liquid chromatographic determination of gypsogenin 3,O-glucuronide. J Chromatogr A 477: 413-419.

Henry M, Rochd M, Bennini B (1991). Biosynthesis and accumulation of saponins in *Gypsophila paniculata*. Phytochemistry 30: 1819-1821.

Hosokawa K, Minami M, Kawahara K, Nakamura I, Shibata T (2000). Discrimination among three species of medicinal *Scutellaria* plants using RAPD markers. Planta Med 66: 270-272.

Iankulov I (1974). Plant growth and changes in the amount and stability of root foam in Samokov 477 variety of *Gypsophila paniculata* L. during its growing season. Plant Sci. 11: 77-84.

ICH (2005). International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use, Validation of Analytical Procedures: Text and Methodology Q2(R1). Geneva, Switzerland: International Council on Harmonisation.

Korkmaz M, Dogan NY (2015). Biogeographic pattern of genetic diversity detected by RAPD and ISSR analysis in *Gypsophila* (Caryophyllaceae) species from Turkey. Genet Mol Res 14: 8829-8838.

Korkmaz M, Özçelik H (2011). Economic importance of *Gypsophila* L., *Ankyropetalum Fenzl* and *Saponaria* L. (Caryophyllaceae) taxa of Turkey. Afr J Biotechnol 10: 9533-9541.

Kovach W (1999). MVSP – A Multivariate Statistical Package for Windows, Ver. 3.1. Pentraeth, UK: Kovach Computing Services.

Li R, Dao Z (2011). Molecular authentication of the traditional Tibetan medicinal plant, *Meconopsis impedita*. Afr J Biotechnol 10: 7493-7496.

Manole A, Banciu C, Indreica A (2015). Genetic diversity within a newly identified population of *Adenophora liliifolia* (L.) A.DC. in Romania: implications for conservation. Ann For Res 58: 1.

Martínez-Nieto MI, Segarra-Moragues JG, Merlo E, Martínez-Hernández F, Mota JF (2013). Genetic diversity, genetic structure and phylogeography of the Iberian endemic *Gypsophila struthium* (Caryophyllaceae) as revealed by AFLP and plastid DNA sequences: connecting habitat fragmentation and diversification. Biol J Linn Soc 173: 654-675.

Mohan M, Nair S, Bhagwat A, Krishna TG, Yano M, Bhatia CR, Sasaki T (1997). Genome mapping, molecular markers and marker-assisted selection in crop plants. Mol Breeding 3: 87-103.

Nei M, Li WH (1979). Mathematical model for studying genetic variation in terms of restriction endonucleases. P Natl Acad Sci USA 76: 5269-5273.

Okoń S, Paczos-Grzęda E, Łoboda M, Sugier D (2014). Identification of genetic diversity among *Arnica montana* L. genotypes using RAPD markers. Acta Sci Pol-Hortoru 13: 63-71.

Okoń S, Surmacz-Magdziak A, Paczos-Grzęda E (2013). Genetic diversity among cultivated and wild chamomile germplasm based on ISSR analysis. Acta Sci Pol-Hortoru 12: 43-50.

Prevost A, Wilkinson MJ (1999). A new system of comparing PCR primers applied to ISSR fingerprinting of potato cultivars. Theor Appl Genet 98: 107-112.

Samal KC, Jena RC, Swain SS, Das BK, Chand PK (2012). Evaluation of genetic diversity among commercial cultivars, hybrids and local mango (*Mangifera indica* L.) genotypes of India using cumulative RAPD and ISSR markers. Euphytica 185: 195-213.

Santos E, Matos M, Silva P, Figueiras AM, Benito C, Pinto-Carnide O (2016). Molecular diversity and genetic relationships in Secale. Jpn J Genet 95: 273-81.

Saxonov SV, Senator SA, Rozenberg GS (2011). Conservation issues floristic diversity Volga Basin reference to the Red Book. Izvestiya Samarskogo Nauchnogo Centra RAN 13: 91-100.

Schulman AH (2007). Molecular markers to assess genetic diversity. Euphytica 158: 313-321.

Soodabeh S, Hoda F, Pantea S, Fahimeh M, Gholamreza A (2011). Genetic diversity and relationships detected by inter simple sequence repeat (ISSR) and randomly amplified polymorphic DNA (RAPD) analysis among *Polygonum* species growing in North of Iran. Afr J Biotechnol 10: 18981-18986.

Williams JG, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV (1990). DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res 18: 6531-6535.

Yao S, Ma L, Luo JG, Wang JS, Kong LY (2010). New triterpenoid saponins from the roots of *Gypsophila paniculata* L. Helv Chim Acta 93: 361-374.

Ziętkiewicz E, Rafalski A, Labuda D (1994). Genome fingerprinting by simple sequence repeat (SSR)-anchored polymerase chain reaction amplification. Genomics 20: 176-183.

Table S1. Precision, mean recovery, LOD, and LOQ of gypsogenin 3-*O*-glucuronide (G3G).

Precision			Recovery (%)	LOD (μg/mL)	LOQ (μg/mL)
	Repeatability over 1 day	Reproducibility over 10 days (mean per day)		2.7 LOD = 3.3σ/S or 2.73 × 3.3 3355	8.2 LOD = 10σ/S or 2.73 × 10 3355
Peak area	191.20	191.49	100.50		
	188.91	188.99	99.00		
	194.36	194.36	92.80		
	191.50	187.40	Mean 97.43		
	194.70	194.60	SD 4.08		
	188.0	198.20			
		193.74			
		198.07			
		196.72			
		197.92			
Mean area	191.45	194.15			
SD	2.74	3.83			
RSD (%)	1.43	1.97			

Table S2. Measured traits in the underground parts of 7 *Gypsophila* species (mean from 2015).

Species	*G. scorzonerifolia*	*G. pacifica*	*G. paniculata*	*G. altissima*	*G. acutifolia*	*G. zhegulensis*	*G. oldhamiana*
Lateral roots number (units/plant)							
Min.	10.0	10.0	9.0	9.0	15.0	9.0	13.0
Max.	18.0	16.0	12.0	11.0	17.0	11.0	15.0
Mean ± SD[a]	13.2 ± 2.9	12.2 ± 2.4	10.2 ± 1.2	10.0 ± 0.9	16.0 ± 0.9	10.5 ± 0.8	14.0 ± 0.6
V%[b]	22.2	19.7	11.5	8.9	5.6	8.3	4.5
Root diameter (mm)							
Min.	48.24	51.73	43.21	76.21	70.24	60.21	55.45
Max.	71.24	76.37	60.04	107.3	92.73	80.82	73.17
Mean ± SD[a]	61.44 ± 7.47	67.11 ± 7.03	50.41 ± 6.54	88.75 ± 10.38	81.56 ± 7.25	71.54 ± 6.65	64.72 ± 5.69
V%[b]	12.16	10.48	13.96	11.69	8.89	9.29	8.80
Length of root (cm)							
Min.	36.0	29.0	35.0	28.0	24.0	24.0	25.0
Max.	39.0	35.0	50.0	36.0	38.0	45.0	42.0
Mean ± SD[a]	37.6 ± 1	31.8 ± 2.3	42.4 ± 5.4	30.8 ± 2.9	31.6 ± 4.6	34.4 ± 6.8	32.4 ± 6.02
V%[b]	2.7	7.3	12.8	9.5	14.5	20.0	18.58
Fresh weight of single root (g/plant)							
Min.	665.60	483.40	282.80	436.80	560.50	200.00	374.70
Max.	1123.70	1060.70	709.20	1053.90	948.00	623.30	480.60
Mean ± SD[a]	902.92 ± 183.6	767.02 ± 210.2	425.14 ± 158.7	664.06 ± 220.4	714.90 ± 137.3	407.18 ± 141.3	434 ± 38.8
V%[b]	20.33	27.40	37.32	33.19	19.20	34.75	8.96
Air-dried mass of single root (g/plant)							
Min.	202.90	143.60	94.80	101.50	194.80	73.30	98.10
Max.	418.60	399.30	217.40	370.00	285.00	210.40	137.20
Mean ± SD[a]	311.46 ± 83.4	274.90 ± 93.3	137.26 ± 48.2	215.58 ± 99.6	243.00 ± 35.5	131.65 ± 45.5	117.26 ± 12.5
V%[b]	26.78	33.94	35.12	46.19	14.62	34.57	10.71

[a]SD- Standard deviation, [b]V- variation coefficient.

Table S3. Measured traits in the aboveground parts of 7 *Gypsophila* species (mean from 2015).

Species	*G. scorzonerifolia*	*G. pacifica*	*G. paniculata*	*G. altissima*	*G. acutifolia*	*G. zhegulensis*	*G. oldhamiana*
Height of plant (cm)							
Min.	167.0	146.0	67.0	157.0	158.0	130.0	70.0
Max.	175.0	167.0	86.0	185.0	180.0	137.0	127.0
Mean ± SD[a]	170.8 ± 3	156.6 ± 7.6	74.0 ± 6.5	167.4 ± 11.1	170.4 ± 7.2	133.3 ± 2.3	104.8 ± 20.8
V%[b]	1.8	4.8	8.8	6.6	4.2	1.7	19.7
Number of stems per plant (units/plant)							
Min.	11.0	14.0	9.0	6.0	12.0	21.0	11.0
Max.	27.0	26.0	12.0	32.0	33.0	33.0	16.0
Mean ± SD[a]	18.6 ± 6.2	20.6 ± 4.5	10.2 ± 1	20.6 ± 9.7	19.6 ± 7.7	26.5 ± 4.3	13.6 ± 1.8
V%[b]	33.1	21.6	9.6	46.9	39.5	16.1	53.7
Stem diameter (mm)							
Min.	10.59	7.46	5.48	9.18	7.50	9.95	5.27
Max.	12.39	12.05	9.07	58.13	11.10	12.40	7.17
Mean ± SD[a]	11.46 ± 0.76	9.70 ± 1.77	7.68 ± 1.2	20.26 ± 18.99	9.36 ± 1.31	10.96 ± 0.89	6.41 ± 0.62
V%[b]	6.67	18.28	15.65	93.71	13.98	8.13	9.72
Number of branches (units/stem)							
Min.	15.6	11.6	25.6	12.0	10.4	18.4	9.0
Max.	19.2	16.0	32.8	16.4	16.0	21.2	12.6
Mean ± SD[a]	18.2 ± 1.4	13.8 ± 1.5	28.7 ± 2.9	14.3 ± 1.8	13.5 ± 2	20.1 ± 1.1	10.4 ± 1.3
V%[b]	7.4	10.8	10.1	12.3	14.7	5.2	12.4
Length of a single leaf (cm)							
Min.	10.8	10.9	4.0	11.1	10.1	9.4	5.4
Max.	15.9	13.7	4.8	13.4	12.5	10.4	12.5
Mean ± SD[a]	12.6 ± 1.9	12.8 ± 1	4.6 ± 0.3	12.1 ± 0.9	11.4 ± 0.8	9.8 ± 0.3	8.4 ± 2.5
V%[b]	15.3	7.9	7.0	7.8	6.9	3.4	30.0
Width of a single leaf (cm)							
Min.	2.6	3.0	0.6	3.7	3.4	2.9	1.0
Max.	4.5	4.3	0.9	4.3	4.3	3.8	2.1
Mean ± SD[a]	3.8 ± 0.7	3.8 ± 0.5	0.8 ± 0.1	4.1 ± 0.2	3.7 ± 0.3	3.4 ± 0.3	1.5 ± 0.4
V%[b]	19.6	12.7	12.6	4.9	9.0	9.3	25.6
Fresh weight of aboveground parts of a single plant (g/plant)							
Min.	1336.00	736.00	224.00	776.60	651.80	2268.00	408.70
Max.	3250.00	2097.00	668.00	3030.00	2547.00	4500.00	481.30
Mean ± SD[a]	2396.6 ± 653	1384.80 ± 449.7	352.48 ± 163.2	1948.12 ± 716	1472.62 ± 658.6	2957 ± 802.53	439.62 ± 26.21
V%[b]	22.27	32.47	46.30	36.76	44.72	27.14	5.36
Air-dried mass of aboveground parts of a single plant (g/plant)							
Min.	534.00	317.70	120.00	302.10	244.00	910.90	99.50
Max.	1324.10	985.20	377.00	1285.00	992.80	1997.10	233.40
Mean ± SD[a]	973.52 ± 255.99	592.02 ± 227.6	214.82 ± 91.02	761.26 ± 313.03	573.9 ± 243.7	1259.3 ± 385	177.54 ± 48.7
V%[b]	26.29	38.45	42.37	41.12	42.46	30.58	27.45

[a]SD- Standard deviation, [b]V- variation coefficient.