Metabolic and chemotaxonomical studies in some *Geum* (*Rosaceae*) species

Strahil Berkov[1], Nedjalka Kasabova[2], Dolja Pavlova[2] & Spassimir Tonkov[2]

[1] Institute of Biodiversity & Ecosystem Research, Bulgarian Academy of Sciences, 23 Acad. G. Bonchev str., 1113 Sofia, Bulgaria, e-mail: berkov_str@yahoo.com (corresponding author)

[2] Sofia University St. „Kl. Ohridski", Faculty of Biology, Department of Botany, 8 Dragan Tsankov blvd., 1164 Sofia, Bulgaria

Received: September 19, 2016 ▷ Accepted: March 31, 2017

Abstract. The metabolite patterns of *Geum reptans*, *G. montanum*, *G. bulgaricum*, and *G. bulgaricum* × *G. reptans* were studied by GC-MS and NIRS. More than 120 compounds have been detected, including many bioactive phenolic compounds, such as arbutine, hydroxycinnamic acid, gallic aicd, caffeic acid, etc. PCA of all GC-MS metabolite patterns resulted in discrimination of *G. reptans*, *G. montanum* and *G. bulgaricum*, but not of *G. bulgaricum* × *G. reptans* from *G. bulgaricum*, which was achieved by PCA of NIRS metabolite patterns. PCA of the metabolite patterns showed that the studied metabolite groups (phenolic acids, organic acids, saccharides, glycerides, fatty acids, fatty alcohols, hydrocarbons, tocoferols and phytosterols) have diverged at different level during speciation and, therefore, they have different chemotaxonomical values.

Key words: chemotaxonomy, *Geum*, GC-MS, metabolomics, NIRS

Introduction

Metabolites are the end products of cellular regulatory processes and their levels can be regarded as the ultimate response of biological systems to genetic or environmental changes (Fiehn 2002). High throughput metabolomics platforms (MS- and NMR-based) allow rapid identification of a large number of metabolites (Hollywood & al. 2006), screening for secondary and bioactive metabolites in plants, and searching for new natural products, which would save time and money (Torras-Claveria & al. 2010; Berkov & al. 2011). NMR, GC-MS and LC-MS metabolite profiling have been used for classification of plants and microorganisms. Global NMR, LC or GC metabolite patterns have been processed for classification of the samples, e.g. species in recent metabolomics studies (Kim & al. 2010; Farag & al. 2012; Plischke & al. 2012). Specific metabolite groups such as terpenes, alkaloids, flavonoids have been largely employed in classical chemotaxonomic studies (Aliferis & al. 2013; Tundis & al. 2014, etc.).

Although the genus *Geum* L. consists of about 50 species, only a few of them have been studied for their chemical composition. Some are known in folk medicine for their anti-inflammatory, astringent, diuretic, and antimicrobial properties (Panizzi & al. 2000; Cheng & al. 2011). Fractions and isolated compounds from *Geum* species have shown activity against HIV-I protease, herpes simplex viruses HSV-1 and HSV-2, and *Helicobacter pylori* (Xu & al. 1996, 2000; Kurokawaru & al. 1998). Their extracts have been found to contain tannins, flavonoids, phenolics, and triterpenoids, while the essential oils obtained from roots are rich in monoterpenoides and fatty acids. The most studied species are *Geum japonicum* Thunberg, *G. rivale* L. and *G. urbanum* L. (Schultze & Wollman 1995; Cheng & al. 2011; Owczarek & Gudej 2013; Owczarek & al. 2013).

Eight *Geum* species grow in Bulgaria (Assenov 1973) and some limited data are available on the bioactive constituents of a few of them. The composition of essential oils from *G. reptans* L., *G. bulgaricum* Panc. and *G. bulgaricum* × G. *rivale* (Schultze & Wollman 1995; Wollman & Schultze 1995) has been reported. The species *G. bulgaricum* has shown several compounds, including flavonoids and phenolic acids (Kaminska & Assenov 1971). Methyl-*b*-D-glucopyranoside was found as one of the main soluble saccharides in *G. montanum* (Aubert & al. 2004).

The species *G. reptans*, *G. montanum* L., *G. bulgaricum*, and the hybrid *G. bulgaricum* × *G. reptans* distributed in the Rila Mts (Bulgaria) belong to the subgenus *Oreogeum* of genus *Geum*, while *G. reptans* and *G. montanum* are distributed in the mountains of Central and South Europe. *Geum bulgaricum* is a relict and endemic species for the Balkan Peninsula and in Bulgaria it is locally distributed in the Rila Mts. This species is distinct in comparison to the other two species. It was described from subalpine shrub communities of *Pinus mugo* Turra and *Juniperus sibirica* Burgsd., and in the subalpine grasslands on silicate terrains at 2000–2800 m a.s.l. (Assenov 1973).

The aim of the present study was to investigate the metabolite profiles of the three closely related species *G. reptans*, *G. montanum*, *G. bulgaricum*, and the hybrid *G. bulgaricum* × *G. reptans*, searching for bioactive molecules, on the one hand, and to evaluate the chemotaxonomical value of the different metabolite groups, on the other. Chemical divergence in plants reflects their genetic evolution and, similar to morphological divergence, some metabolite groups may provide more and other less taxonomical and evolutionary information. To the best of our knowledge such a metabolomics-based chemotaxonomical approach is applied for the first time.

Material and methods

Plant material

Aerial parts from 4–5 individuals of *G. reptans*, *G. montanum*, *G. bulgaricum* and *G. bulgaricum* × *G. reptans* were randomly collected at flowering stage in July 2012. The populations of all species were sympatric at 2300 m a.s.l. in the Rila Mts., Seven Rila Lakes locality (N 42° 11'.55 13"; E 23° 19'.08 28"). The samples were lyophilized and then powdered in a mill. Voucher specimens of the species were deposited at the Herbarium of Sofia University (SO).

Extraction

A multi-step sample preparation protocol has been adapted, following Schielmann & al. (2008). Samples from five individuals per each species were analyzed. Fifty milligrams of plant material were placed in 2 ml Ependorf tubes and extracted with 1 ml of MeOH for 30 min at 70 °C after addition of internal standards of 50 μg of nonadecanoic acid for apolar, 50 μg of ribitol for polar and 3,4 dichloro-4-hidroxy benzoic acid for phenolic compounds.

Then the extract was centrifuged and an aliquot of 800 μl was transferred to other Ependorf tubes. After addition of 500 μl H_2O and 500 μl of $CHCl_3$ and vortexing for 2 min, the mixture was centrifuged and the chloroform fraction was separated, evaporated and transmethylated with 2 % of H_2SO_4 in MeOH at 90 °C for 2 h. The lipids were extracted with *n*-hexane (2 × 500 μl), which was dried with anhydrous Na_2SO_4 and evaporated to obtain lipid fractions. An aliqote of 50 μl from the aqueous fraction was placed in a glass vial and evaporated in a speed-vac to obtain fractions with polar compounds (aminoacids, saccharides, organic acids, and free phenolic acids). The rest of the aqueous fraction was hydrolyzed with 0.5 ml of 1N NaOH for 12h at 80 °C to obtain a phenolic fraction. After acidification to pH 1–2 with conc. HCl, the phenolic compounds were extracted with EtOAc (2 × 500 μl), which was dried with anhydrous Na_2SO_4 and evaporated.

Lipid fractions were silylated with 50 μL of N,O-bis-(trimethylsilyl)trifluoro-acetamide (BSTFA) in 50 μL of pyridine for 2 h at 50 °C. Polar fractions were first methoxymated with 100 μL of methoxyamine hydrochloride (MOA, 20 mg mL–1 in pyridine) for 90 min at 50 °C to stabilize the carbonyl moiety and prevent ring formation in sugars, then silylated by addition of 100 μL of BSTFA and heating at 50 °C for 1 h. The mixtures were diluted with 200 μL of chloroform for further GC-MS analysis.

Metabolite analysis

GC-MS analyses were performed on a Hewlett Packard 6890$^+$/MSD 5975 instrument (Hewlett Packard, Palo Alto, CA, USA) operating in EI mode at 70 eV. A HP-5 MS column (30 m × 0.25 mm × 0.25 mm) was used. The temperature program was: 100–180 °C at 15 °C × min^{-1},

180–300 at 5 °C × min^{-1} and 10 min hold at 300 °C. The injector temperature was 250 °C. The flow rate of carrier gas (helium) was 0.8 mL × min^{-1}. The split ratio was 1:15. One μL of the solution was injected.

Metabolites which exceeded 0.1 μg per dry sample are listed in Table 1. They were identified as TMSi derivatives comparing their mass spectra and Kovats Indexes (RI) with those of an online available plant specific database (The Golm Metabolome Database; http://csbdb.mpimp-golm.mpg.de/csbdb/gmd/home/gmd_sm.html), the NIST 05 database and mass spectra available in the online lipid library (http://www.lipidlibrary.co.uk/ms/ms01/index.htm), NIST 05 database and literature data, as indicated in Table 1. The measured mass spectra were deconvoluted by the Automated Mass Spectral Deconvolution and Identification System (AMDIS) before comparison with the databases. Then the spectra of the individual components were transferred to the NIST Mass Spectral Search Program MS Search 2.0, where they were matched against reference compounds of the NIST Mass Spectral Library 2005 and the Golm Metabolome Database. RI of the compounds were recorded with standard *n*-hydrocarbon calibration mixture (C9-C36) (Restek, Cat no. 31614, supplied by Teknokroma, Spain) using AMDIS 3.6 software.

Table 1. Metabolites in *G. reptans*, *G. montanum*, *G. bulgaricum*, and *G. bulgaricum* × *G. reptans*.
The values ± SD (μg) represent the response ratios calculated for each analyte relative to the internal standard.

Compounds	*RI*	*G. reptans*	*G. montanum*	*G. bulgaricum*	*G. hybrid*
Hydrocarbones					
Tetracosane (**1**)	1401	5,0±0,2	5,1±1,1	0,6±0,1	0,6±0,1
Hydrocarbone, branched 1 (**2**)	1490	4,6±0,8	4,3±0,4	1,0±0,2	1,1±0,1
Pentacosane (**3**)	1500	0,3±0,1	0,3±0,1	trace	trace
Hydrocarbone, branched 2 (**4**)	1535	3,6±0,7	3,1±0,3	0,7±0,2	0,9±0,1
Hexacosane (**5**)	1598	4,5±0,4	4,3±0,9	1,9±0,3	2,2±0,4
Heptacosane (**6**)	1702	0,1±0,1	0,2±0,3	traces	traces
Hydrocarbone, branched 3 (**7**)	1707	4,8±0,8	4,5±0,4	1,7±0,2	2,1±0,2
Hydrocarbone, branched 4 (**8**)	1751	3,1±0,7	2,8±0,2	1,1±0,2	1,4±0,2
Hydrocarbone, branched 5 (**9**)	1762	traces	0,2±0,0	0,6±0,2	0,9±0,3
Octadecane (**10**)	1800	3,0±0,5	2,6±0,6	1,9±0,3	2,5±0,4
Hydrocarbone, branched 6 (**11**)	1886	0,2±0,1	0,3±0,0	0,1±0,1	0,5±0,2
Nonadecane (**12**)	1897	1,0±0,2	1,6±0,9	0,5±0,1	0,6±0,1
Hydrocarbone, branched 7 (**13**)	1955	0,4±0,1	0,5±0,0	0,1±0,0	0,1±0,1
Hydrocarbone, branched 8 (**14**)	1967	2,2±0,5	2,3±0,3	0,8±0,1	1,3±0,3
Eicosane (**15**)	2000	2,7±0,4	2,4±0,4	1,7±0,3	2,2±0,5
Hydrocarbone, branched 9 (**16**)	2140	3,2±0,7	3,2±0,6		
Docosane (**17**)	2198	1,2±0,4	0,9±0,3	1,0±0,3	1,4±0,4
Untriacontane (**18**)	3093	0,7±0,7	0,2±0,2	0,2±0,1	0,3±0,2
		59,8±11,2	**57,1±10,8**	**23,6±4,6**	**25,9±10,1**
Fatty acids					
Tetradecanoic acid (14:0) (**19**)	1723	2,0±1,6	1,7±1,0	1,5±0,4	0,9±0,2
Pentadecanoic acid (15:0) (**20**)	1827	1,5±1,5	0,4±0,4	0,2±0,0	0,2±0,1
Metylpentadecanoic acid (16:1) (**21**)	1921	7,8±1,3	6,8±2,4	3,4±1,1	5,0±3,0
Hexadecanoic acid (16:1) (**22**)	1927	193,8±14,6	141,6±33,3	70,6±6,7	89,6±30,1
Heptadecanoic acid (17:0) (**23**)	2028	2,6±0,6	3,9±2,8	1,0±0,3	1,0±0,4
Octadecadienoic acid (C18:2) (**24**)	2099	72,7±29,9	118,1±17,3	64,8±22,9	49,2±18,0
Octadecatrienoic acid (C18:3) (**25**)	2112	756,0±66,5	479,2±59,4	230,1±33,4	295,6±89,8
Octadecanoic acid (18:0) (**26**)	2129	94,5±28,3	86,9±27,7	53,8±23,0	58,5±18,8
Eicosanoic acid (20:0) (**27**)	2328	trace	4,3±2,6	3,0±0,8	45,0±40,1
Docosenoic acid (C22:1) (**28**)	2503	18,9±9,1	11,1±10,9	5,6±5,7	7,4±7
Docosanoic acid (22:0) (**29**)	2527	10,1±4,6	1,6±0,6	2,1±0,3	2,0±0,9
Tricosanoic acid (23:0) (**30**)	2627	0,4±0,2	0,2±0,1	0,20±0,1	1,4±1,6
Tetracosanoic acid (24:0) (**31**)	2727	21,8±3,8	3,8±2,8	3,4±1,7	4,4±2,4
Hexacosanoic acid (26:0) (**32**)	2927	3,3±1,2	0,2±0,2	0,5±0,2	0,5±0,1
		1185,4±163,1	**859,8±161,5**	**440,2±96,6**	**560,7±212,5**

Table 1. Continuation.

Compounds	*RI*	*G. reptans*	*G. montanum*	*G. bulgaricum*	*G. hybrid*
Hydroxyfatty acids					
3-Hydroxyhexadecanoic acid (**33**)	2130	14,7±4,4	16,8±12,0	53,1±27,7	9,8±2,4
3-Hydroxyoctadecanoic acid (**34**)	2331	873,9±87	31,923,9±	118,4±52	69,2±34,3
3-Hydroxyeicosnoic acid (**35**)	2520	8,1±3,6	0,2±0,2	1,7±0,6	1,2±0,8
3-Hydroxydocosnoic acid (**36**)	2716	4,8±1,6	0,6±0,4	1,6±0,5	1,3±0,7
2-Hydroxytetracosanoic acid (**37**)	2913	2,8±0,3	1,7±1,2	0,7±0,1	0,9±0,4
2-Hydroxypentacosanoic acid (**38**)	3010	0,1±0,1	0,2±0,2	0,1±0,1	0,1±0,0
		899,6±95,3	**50,8±37,5**	**174,1±80,6**	**81,1±37,9**
Fatty alcohols					
1-Hexadecanol (**39**)	1961	19,3±12,3	1,2±0,3	14,5±4,5	20,9±7
1-Octadec-9Z-enol (**40**)	2138			35,3±19,4	19,5±21,9
1-Octadecanol (**41**)	2158	4,3±1	4,2±1,2	12,5±4,6	16,5±5,4
1-Tetracosanol (**42**)	2746	24,3±4,8	2,7±0,6	16,7±6,6	6,1±3,7
1-Hexacosanol (**43**)	2943	46,4±10,9	9,3±2,7	10,8±3,1	10,2±3,8
		109,5±29	**39,4±9,7**	**100,4±39,8**	**86,7±46,2**
Tocoferols					
gamma-Tokoferol (**44**)	3116		0,6±0,6		
alpha-Tocopherol (**45**)	3152	7,2±4,9	4,3±4,0	1,7±3	1,3±1
		7,2±4,9	**5,0±4,6**	**1,7±3**	**1,3±1**
Glycerides					
Glycerol	1289	52,0±11,0	35,9±6,3	21,3±3,1	25,6±5,9
1-Monohexadecanoylglycerol (**46**)	2599			12,1±4,6	16,4±5,1
1-Monooctadecanoylglycerol (**47**)	2792			11,6±4,3	16,9±1,8
Unidentified glyceride (**48**)	2814	0,2±0,0	0,2±0,2		0,1±0,1
		52,3±11,0	**36,1±6,5**	**45,0±12,1**	**59,0±12,9**
Phytosterols					
Sitosterol acetate (**49**)	3101	1,2±0,8	0,9±0,5	0,4±0,2	0,3±0,3
Cholesterol (**50**)	3161	2,6±2,3	1,5±0,1	0,5±0,5	1,0±0,2
Campestrol (**51**)	3263	4,6±0,9	4,2±0,6	1,6±0,4	2,4±0,5
beta-Sitosterol (**52**)	3359	148,6±24,7	123,8±13,0	52,1±10,1	60,7±15,6
Unidentified sterol (**53**)	3401	5,0±1,6	0,9±0,3	1,2±0,8	0,4±0,6
		162,0±30,3	**131,2±14,5**	**55,9±12,0**	**64,7±17,2**
Polyenes					
Phytol 1 (**54**)	1847	3,4±0,6	2,0±0,4	0,4±0,1	0,7±0,7
Phytol 2 (**55**)	1863	6,30,8±	5,2±1,1	2,0±0,4	3,4±1,1
Phytol 3 (**56**)	1868	6,4±1,1	4,1±0,8	0,8±0,2	1,3±0,5
Phytol 4 (**57**)	1881	5,8±0,9	4,3±0,8	1,4±0,3	2,5±0,9
Phytol 5 (**58**)	1904	9,5±1,7	6,8±1,1	1,9±0,2	3,0±0,8
Phytol 6 (**59**)	2015	4,0±0,8	3,9±3,4	1,0±0,2	1,5±0,5
Phytol 7 (**60**)	2037	10,1±1,5	7,3±1,5	2,3±0,7	3,9±1,8
Phytol 8 (**61**)	2068	3,5±3,0	6,7±1,7	3,4±0,8	5,1±1,8
		49,1±10,4	**40,2±10,7**	**13,3±2,8**	**21,4±8,1**
Unknown lipids					
UC-geum 1 (**62**)	2181	14,6±2,9	10,1±1,9	3,8±0,5	5,6±2,0
UC-geum 2 (**63**)	2474	8,2±6,7			
UC-geum 3 (**64**)	2513	4,9±3,4			
UC-geum 4 (**65**)	2558	8,6±6,2			
UC-geum 5 (**66**)	2574	4,4±2,4			
UC-geum 6 (**67**)	2579	2,9±4,6			
UC-geum 7 (**68**)	2596	4,9±5,8			
UC-geum 8 (**69**)	3518	1,6±1,4	4,8±3,4	56,3±31,4	31,0±27,2

Table 1. Continuation.

Compounds	*RI*	*G. reptans*	*G. montanum*	*G. bulgaricum*	*G. hybrid*
UC-geum 9 (**70**)	3684	41,6±6,6	41,6±4,4	22,8±2,9	19,6±11,1
		91,6±39,9	**56,6±9,7**	**83,0±34,8**	**56,2±40,4**
Organic acids					
Succinic acid (**71**)	1325	8,0±2,5	10,6±2,6	3,7±0,6	5,8±5,0
Glyceric acid (**72**)	1346	1,1±0,5	0,7±0,3	0,7±0,4	1,2±0,9
Aspartic acid (**73**)	1479	1,3±0,6	0,5±0,3	0,2±0,1	0,6±0,3
Malic acid (**74**)	1491	99,2±22,7	22,6±8,5	232,1±65,9	188,3±115,1
Butanoic acid (**75**)	1533	2,8±2,8	2,3±1,2	4,2±1,8	10,4±5,2
Threonic acid (**76**)	1555	1,5±0,4	0,2±0,1	0,7±0,1	1,0±0,2
2,3,4,5-Tetrahydroxypentanoic acid-1,4-lactone (**77**)	1666	2,8±0,5	2,8±0,6	1,4±0,2	1,5±1,0
Ribonic acid (**78**)	1794	0,5±0,2	0,2±0,1	0,6±0,1	0,7±0,2
Shcimic acid (**79**)	1826	20,7±9,5	5,8±4,1	2,9±3,5	11,7±15,4
		138,1±39,6	**45,7±17,8**	**246,6±72,7**	**221,2±143,3**
Polyoles					
Erythritol (**80**)	1515	2,1±0,2	0,5±0,3	0,5±0,1	0,8±0,2
Arabitol (**81**)	1742	1,6±1,1	0,8±0,3	0,9±0,4	0,5±0,2
Myo-Inositol (**82**)	2127	273,6±85	318,1±113,4	235,3±44,1	298,5±169,3
		277,4±86,3	**319,5±113,9**	**236,7±44,6**	**299,8±169,7**
Monosaccharides					
Monosaccharide1 (**83**)	1727	59,9±25,9	29,9±7,0	56,6±16,5	62,2±27,8
Monosaccharide2 (**84**)	1764	71,2±9,4	58,2±13,1	33,5±8,3	36,2±14,1
Monosaccharide3 (**85**)	1780	1,5±0,5	0,8±0,2	1,8±0,7	1,3±0,6
Monosaccharide4 (**86**)	1785	82,6±37,8	39,7±9,7	66,5±19,9	74,3±32,2
Monosaccharide5 (**87**)	1802			45,9±35,9	50,9±11,8
Monosaccharide6 (**88**)	1818			65,1±51,7	44,9±12,8
Fructose 1 (**89**)	1834	659,3±302,7	439,7±41	329,2±96,5	380,7±58,4
Fructose 2 (**90**)	1842	690,2±39,8	628,6±144,1	389,8±51,8	385,4±57,6
Fructose 3 (**91**)	1850	690,6±142,5	537,1±91,5	284,1±116,4	313,0±54,3
Glucose 1 (**92**)	1926	1275±125,9	1045,2±137,8	673,3±95,4	381,0±380,4
Glucose 2 (**93**)	1930	691,4±120,6	744,6±223,2	342,2±160,8	329,5±122,6
Unidentified monosaccharide (**94**)	1950	181,1±184,4	85,0±43,0	5,7±3,3	25,8±31,2
Unidentified monosaccharide (**95**)	2018	1166,6±248,0	1044,4±237,5	629,3±186	705,2±97,6
Galactosylglycerol (**96**)	2361	6,8±3	4,3±2,1	4,0±2,1	3,9±1,0
		5576,2±1240,6	**4657,4±950,1**	**2927±845,4**	**2794,2±902,3**
Disaccharides					
Disaccharide 1 (**97**)	2560	17,1±16,2	5,9±1,5	2,8±2	3,5±2,1
Disaccharide 2 (**98**)	2573	37,6±17,5			2,3±1,6
Disaccharide 3 (**99**)	2627	18,8±12,5			0,8±1,3
Sucrose (**100**)	2706	1017,6±233,8	512,0±138,9	468,2±118,6	379,1±142,0
Threhalose (**101**)	2819	49,4±84,7	1,3±0,5	1,4±1,3	2,4±1,9
Disaccharide 4 (**102**)	2890	25,9±9,1			10,5±7,5
Disaccharide 5 (**103**)	2898	328,2±325,0			22,5±26,5
		1494,6±698,8	**519,2±140,9**	**472,4±121,9**	**421,0±184,9**
Trisaccharides					
Trisaccharide 1 (**104**)	3209	20,7±6,2	12,1±3,4	9,7±2,1	12,7±4,8
Trisaccharide 2 (**105**)	3215			13,5±5,9	2,7±4,6
Trisaccharide 3 (**106**)	3253			16,9±12,9	5,0±5,8
Trisaccharide 3 (**107**)	3335		16,7±7,4	27,8±24,3	10,0±8,4
Rafinnose (**108**)	3513	65,9±44,5	28,0±18,4	90,9±61,5	47,5±17,7
		86,7±50,7	**56,8±29,2**	**158,8±106,6**	**77,8±41,2**

Table 1. Continuation.

Compounds	*RI*	*G. reptans*	*G. montanum*	*G. bulgaricum*	*G. hybrid*
Phenolic compounds					
Arbutine (**109**)	2629		0,4±0,2	0,1±0,1	trace
Salicilic acid (**110**)			trace	trace	trace
Apocynin, 4'-Hydroxy-3'-methoxyacetophenone (**111**)	1630				
4(p)-Hydroxybenzoic acid (**112**)	1633	7,0±6,1	6,5±4,7	178,7±321,7	322,8±461,7
3,5-Dihydroxybenzoic acid (**113**)	1731			249,3±433	170,4±204,6
Dihydrocinnamic (**114**)	1769		1,5±0,3		
Phenolic compound 1 (**115**)	1730	14,1±14,5			
Vanilic acid (**116**)	1773	40,7±29,7	4,1±1,0		
Phenolic compound 2 (**117**)	1787	74,6±116,7			19,5±25,3
Protocatechuic acid (**118**)	1832	25,5±28,3	32,2±5,9		24,2±48,3
3,4-Dihydroxyphenylacetic acid (**119**)	1843	0,2±0,4	4,6±0,7		
Phenolic compound 3 (**120**)	1893		5,8±1,5		
Phenolic compound 4 (**121**)	1907	5,3±7,8	19,0±14,3	5094,9±1921,2	9039,8±3810,0
trans-Hydroxycinnamic acid (**122**)	1948	28,3±9,5	89,9±28,9	154,9±131,3	282,8±152,8
3,4-Hydroxycinnamic acid methyl ester (**123**)	2022				32,6±28,1
Gallic acid (**124**)	1979	592,8±765,0	1218,9±527,6	140,0±79,1	213,3±107,7
trans-Ferulic acid (**125**)	2103	31,0±14,3	11,7±2,6	6,3±69,6	33,8±23,1
trans-Caffeic acid (**126**)	2150	10,3±3	114,5±14,2	289±160	205,1±80,1
Phenolic compound 5 (**127**)	2235		67,2±15,3		
trans-Sinnapic acid (**128**)	2256	4,2±3,2	5,3±2,1		
		903,8±1105,2	**1581,1±619,1**	**6113,0±3046,3**	**10344,3±4421,5**
Polyphenols					
epi-Catechine (**129**)	2922	28,5±10	580,9±260,2	31,9±25,9	105,3±168,0

NIRS Analysis

About 0.5 g of lyophilized and powdered plant samples were scanned from 1100 to 2498 nm on a NIR System model 6500 M (Foss).

Statistical analysis

The principal component analysis (PCA) of the data matrix was performed with Minitab 15 (Minitab Inc.) to check the overall chemical pattern of the studied samples.

Results and discussion

More than 120 compounds were detected by GC-MS in the studied *Geum* species, including organic acids, phenolic acids, polyphenols, fatty acids, fatty alcohols, hydrocarbons, mono-, di-, and tri-saccharides, polyoles, phytosterols, and traces of triterpenes. Some compounds showed high SD values, probably due to individual variability. Most of the detected compounds in this study are reported for the first time for *G. reptans*, *G. montanum*, *G. bulgaricum*, and *G. bulgaricum* × *G. reptans*, considering the scanty phytochemical data on these species. The species showed a similar general metabolite pattern. The differences recorded were both mainly in the metabolite abundance and in some secondary metabolites. The abundance of lipids, including hydrocarbons, fatty acids, hydroxyl fatty acids, and phytosterols, was about twice higher in *G. reptans* and *G. montanum* as compared to *G. bulgaricum* and *G. bulgaricum* × *G. reptans*. The major compounds in the lipid fraction, as expected, were octadecatrienoic (C18:3), followed by hexadecenoic acid (C16:1). Mention deserves the high abundance of 3-hydroxyoctadecanoic acid (**34**) in *G. reptans* and the presence of 1-octadec-9*Z*-enol (**40**) in *G. bulgaricum* and *G. bulgaricum* × *G. reptans*. A number of unidentified compounds in the lipid fraction were detected in *G. reptans*. Five phytosterols were found, the main being *beta*-sitosterol (**52**). Due to abundance of malic acid, the total organic acids were much higher in *G. bulgaricum* and *G. bulgaricum* × *G. reptans*. The accumulation of monosaccharides was *ca*. twice higher in *G. montanum* and *G. reptans* as compared to the other two

species. The most abundant were fructose and glucose. Considerable amounts of compound **95** were observed in these species. The mass spectra, as well as the *RI* of this metabolite are similar to that of glucose and it may be methyl-d-D-glucopyranoside, a compound found to be one of the major soluble saccharides in the leaves of *G. montanum* (Abuert & al. 2004). The main di- and tri-saccharides were sucrose and raffinose. *Epi*-catechine (**129**), a flavane-3-ol, was detected in the polar fraction in considerable amounts in *G. montanum*. Waste biological activities were reported for this compound, including antioxidant and MAO inhibitory among others (Hou & al. 2005). Traces of arbutine, a hydroquinone glycoside, were detected in the polar fraction of *G. montanum*, *G. bulgaricum* and *G. bulgaricum* × *G. reptans*. This compound acts as a skin-lightening agent inhibiting the formation of melanin (O'Donoghue 2006). Analysis of the phenol fraction after hydrolysis resulted in detection of other 19 phenolic compounds (Fig. 1.). Some of them (**115, 117, 120**, **121** and **127**) were not identified, but their mass spectral fragmentation indicated the presence of a phenolic moiety. *Geum reptans* and *G. montanum* showed similar phenolic acids patterns, like those of *G. bulgaricum* and *G. bulgaricum* × *G. reptans*. The most abundant compounds in the first two species were gallic acid (**124**), compound **117**, apocynin (**111**) in *G. reptans* and *trans*-caffeic acid (**126**) in *G. montanum*. Gallic acid possesses antioxidant and antibacterial properties (Borges & al. 2013) and has been found in other *Geum* species as a major phenolic compound (Kaminska & Assenov 1971; Cheng & al. 2011). As it was not detected in a free form, most probably it originated from the hydrolysable tannins characteristic of *Geum* species (Cheng & al. 2011). Acting as an inhibitor of NADPH oxidase, apocynin (**111**) has anti-asthmatic and anti-atherosclerosis activity (Van den Worm & al. 2001). Being one of the main phenolic acids, caffeic acid (**126**) found also in *G. bulgaricum* and *G. bulgaricum* × *G. reptans* is known for its antioxidant and antimicrobial activities (Gülçin 2006). The major phenolic compounds in *G. bulgaricum* and the hybrid species were compound **121** and caffeic acid.

At first glance on similarity of GC-MS metabolite patterns, two groups of species can be seen in *G. reptans*/*G. montanum* and G. *bulgaricum*/*G. bulgaricum* × *G. reptans*. Inter-species discrimination, however, was achieved with a multivariate analysis.

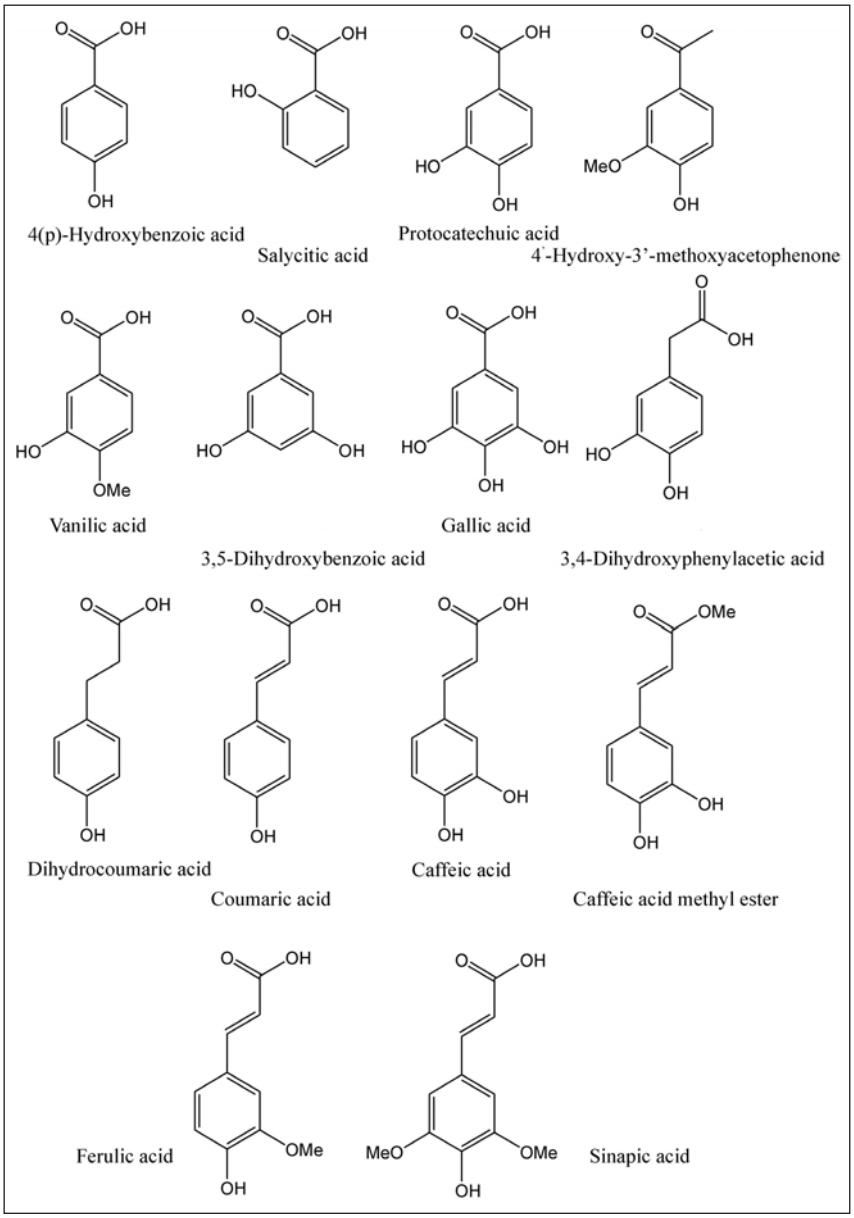

Fig. 1. Phenolic compounds identified in *G. reptans*, *G. montanum*, *G. bulgaricum*, and *G. bulgaricum* × *G. reptans*.

PCA of total GC-MS metabolite patterns resulted in the separation of *G. reptans*, *G. montanum* and G. *bulgaricum*. *G. bulgaricum* × *G. reptans*, however, was not well discriminated from *G. bulgaricum*, which is considered as one of its parents (Fig. 2A). Discrimination of *G. bulgaricum* × *G. reptans* from *G. bulgaricum* was achieved by PCA of NIRS metabolite patterns (Fig. 3) indicating the presence of compounds with chemotaxonomical value that cannot be analyzed by GC-MS, on the one hand, and that NIRS is a powerful approach in chemotaxonomic studies, despite the fact that it does not provide information on individual compounds, on the other.

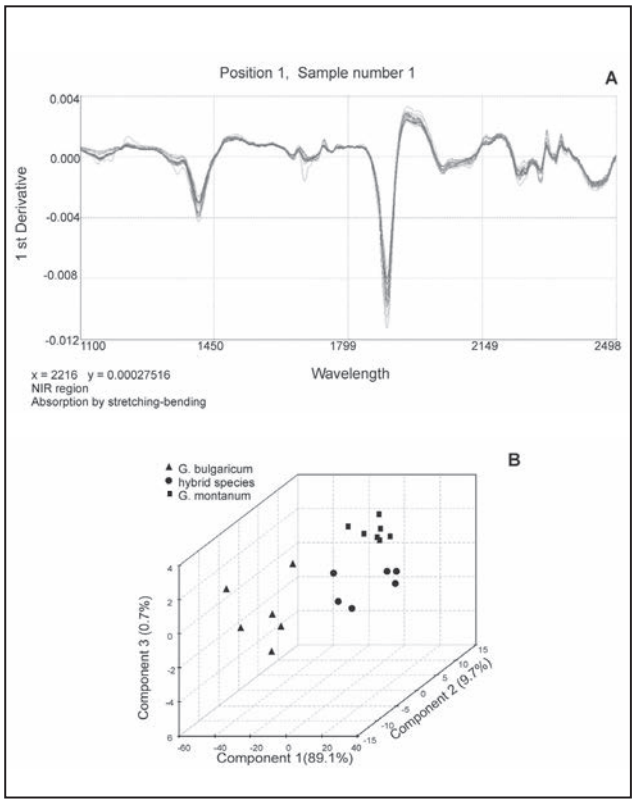

Fig. 3. NIRS spectra (**A**) and PCA (**B**) of *G. montanum* (■), *G. bulgaricum* (▲) and *G. bulgaricum* × *G. reptans* (●).

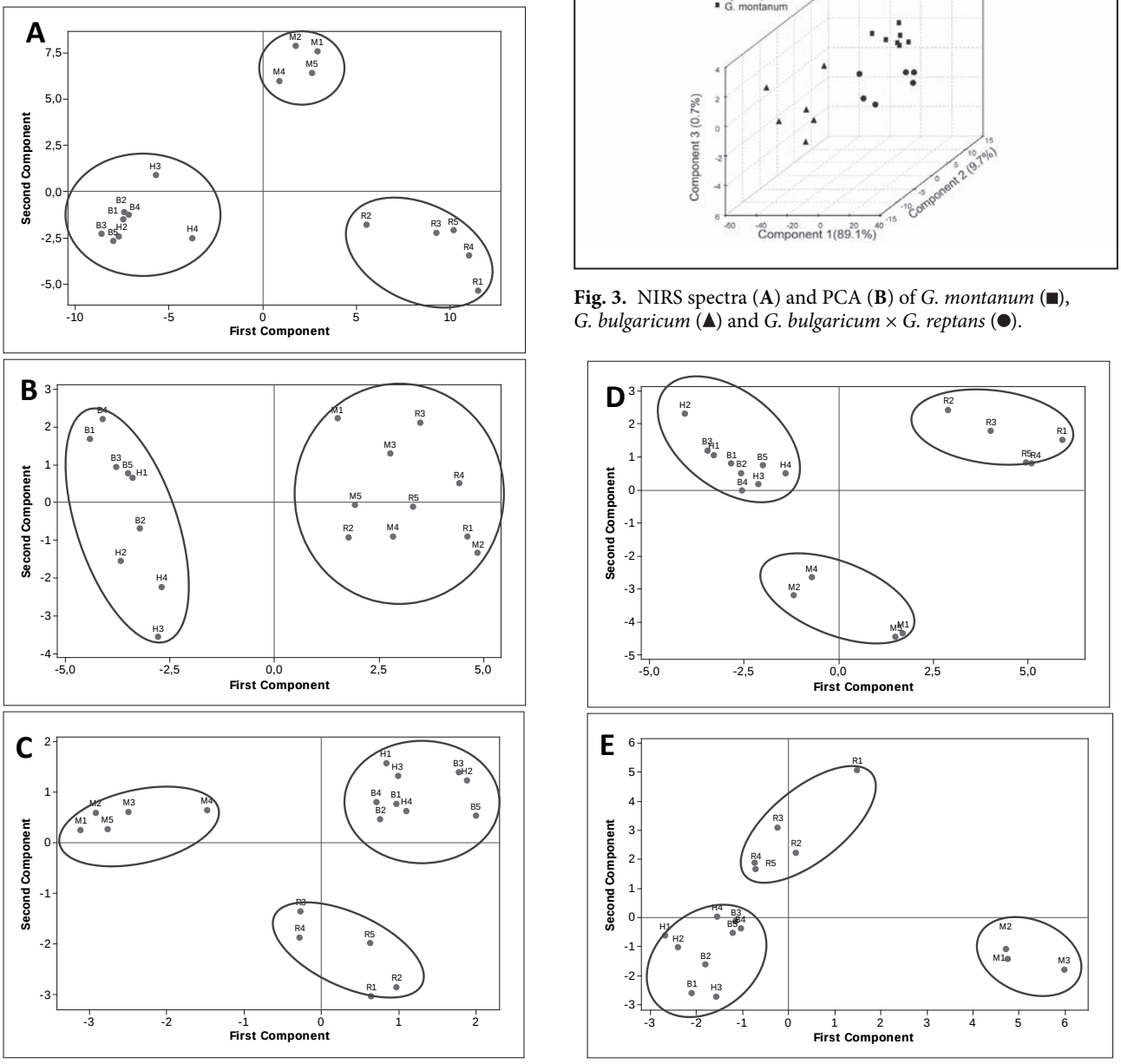

Fig. 2. PCA analysis of metabolites from *G. reptans* (R), *G. montanum* (M), *G. bulgaricum* (B), and *G. bulgaricum* × *G. reptans* (H). **A** – total metabolite pattern, **B** – hydrocarbons, **C**– fatty alcohols, **D** – fatty acids and fatty alcohols, **E** – phenolic compounds.

To evaluate the importance of different metabolite groups for chemotaxonomical differentiation of the studied species, PCA was performed separately for the hydrocarbons, fatty acids, fatty alcohols, organic acids, saccharides, phenolics, and phytosterols (Fig. 2). The results indicated that the patterns of some metabolite groups have diverged more than others, resulting in a higher ability for discrimination of these closely related species. Thus, good intra-species discrimination was obtained by PCA of fatty alcohols, fatty acids and phenolics (Fig. 2C, 2D and 2E). Again, *G. bulgaricum* × *G. reptans* was not well discriminated from *G. bulgaricum*. The PCA of hydrocarbons resulted in poorer discrimination into the two main groups, *G. bulgaricum* with *G. bulgaricum* × *G. reptans*, and *G. reptans* with *G. montanum* (Fig. 2B). Despite that some grouping was observed, the organic acids and saccharides showed a lesser ability for discrimination of the samples. This may be a result of the higher variability of these primary metabolites in daytime, due to the influence of photosynthesis on their levels. Phytosterols seem to form the most conservative group resulting in no discrimination between the samples.

PCA clearly indicated that *G. bulgaricum* is one of the parents of the hybrid. The species *G. reptans* is considered as the other parent, but this cannot be clearly supported by the present study. PCA of total metabolite patterns and fatty acids and fatty alcohols (Fig. 2A, D) indicates that *G. bulgaricum* × *G. reptans* is closer to *G. reptans*, as they are separated from *G. montanum* by the second component, like its parent *G. bulgaricum*. PCA of phenolic compounds and fatty alcohols also groups *G. bulgaricum* × *G. reptans* and *G. bulgaricum* close to *G. reptans*, separating them from *G. montanum* by the first component. On the other hand, PCA of fatty acids indicates relativity to *G. montanum*. Further genetic studies may elucidate the relationship of *G. bulgaricum* × *G. reptans* with the other species.

Conclusions

The species *G. reptans*, *G. montanum*, *G. bulgaricum*, and *G. bulgaricum* × *G. reptans* are rich sources of bioactive phenolic compounds. The metabolite patterns allow chemotaxonomical differentiation of the studied species. The patterns of metabolite groups (*e.g.* hydrocarbons, fatty acids, fatty alcohols, organic acids, saccharides, phenolics, and phytosterols) in these species have diverged at a different level like their morphological features and, therefore, they have different chemotaxonomical value. The metabolite divergence rather affected the abundance of primary metabolites (*e.g.* organic acids, saccharides, fatty acids, fatty alcohols, hydrocarbons) and the pattern of secondary metabolites (*e.g.* phenolic acids) in terms of qualitative differences.

Acknowledgements. The authors are grateful to Dr. I. Badjakov and Mrs. I. Dincheva from AgroBioInstitute, Sofia, for the NIRS analysis.

References

Aliferis, K., Marc, A., Cubeta, M. & Jabaji, S. 2013. Chemotaxonomy of fungi in the *Rhizoctonia solani* species complex performing GC/MS metabolite profiling. – Metabolomics, **9**: 159-169.

Assenov, I. 1973. Genus *Geum* L. – In: **Jordanov, D.** (ed.), Fl. Reipubl. Popularis Bulgaricae. Vol. **5**, pp. 189-205. In Aedibus Acad. Sci. Bulgaricae, Serdicae (in Bulgarian).

Aubert, S., Choler, P., Pratt, J., Douzet, R., Gout, E. & Bligny, R. 2004. Methyl-b-D-glucopyranoside in higher plants: accumulation and intracellular localization in *Geum montanum* L. leaves and in model systems studied by ^{13}C nuclear magnetic resonance. – J. Exp. Bot., **55**: 2179-2189.

Berkov, S., Bastida, J., Viladomat, F. & Codina C. 2011. Development and validation of a GC-MS method for rapid determination of galanthamine in *Leucojum aestivum* and *Narcissus* ssp.: a metabolomic approach. – Talanta, **83**: 1455-1465.

Borges, A., Ferreira, C., Saavedra, M.J. & Simões, M. 2013. Antibacterial activity and mode of action of ferulic and gallic acids against pathogenic bacteria. – Microb. Drug Resist., **19**: 256-265.

Cheng, X.-R., Jin, H.-Z., Qin, J.-J., Fu, J.-J. & Zhang, W.-D. 2011. Chemical constituents of plants from the genus *Geum*. – Chem. Biodiver., **8**: 203-222.

O'Donoghue, J.L. 2006. Hydroquinone and its analogues in dermatology – a risk-benefit view point. – J. Cosm. Dermatol., **5**: 196-203.

Farag, M., Porzel, A. & Wessjohann, L. 2012. Comparative metabolite profiling and fingerprinting of medicinal licorice roots using a multiplex approach of GC–MS, LC–MS and 1D NMR techniques. – Phytochemistry, **76**: 60-67.

Fiehn, O. 2002. Metabolomics – the link between genotypes and phenotypes. – Pl. Molec. Biol., **48**: 155-171.

Gülçin, I. 2006. Antioxidant activity of caffeic acid (3,4-dihydroxycinnamic acid). – Toxicology, **217**: 213-222.

Hollywood, K., Brison, D. & Goodacre, R. 2006. Metabolomics: Current technologies and future trends. – Proteomics, **6**: 4716-4723.

Hou, W.-C., Lin, R.D., Chen, C.T. & Lee, M.H. 2005. Monoamine oxidase B (MAO-B) inhibition by active principles from *Uncaria rhynchophylla*. – J. Ethnopharmacol., **100**: 216-220.

Kaminska, J. & Assenov, I. 1971. Phytochemical investigation of *Geum bulgaricum* Panc. – Acta Polon. Pharm., **28**(2): 204-208.

Kim, H., Khan, S., Wilson, E., Kricun, P., Meissner, A., Goraler, S., Deelder, A., Choi, Y. & Verpoorte, R. 2010. Metabolic classification of South American *Ilex* species by NMR-based metabolomics. – Phytochemistry, **71**: 773-784.

Kurokawaru, M., Purusotam, T., Hozum, B., Nakano, M., Kadota, S., Namba, T., Kawana, T. & Shiraki, K. 1998. Purification and characterization of eugenin as an anti-herpes virus compound from *Geum japonicum* and *Syzygium aromaticum.* – J. Pharmacol. Exp. Therap., **284**: 728-735.

Owczarek, A. & Gudej, J. 2013. Investigation into biological active constituents of *Geum rivale* L. – Acta Polon. Pharm. Drug Res., **70**: 111-114.

Owczarek, A., Gudej, J. & Kicel, A. 2013. Composition of essential oil from aerial and underground parts of *Geum rivale* and *G. urbanum* growing in Poland. – Nat. Prod. Comm., **8**: 505-508.

Panizzi, L., Catalano, S., Miarelli, C., Cioni, P.L. & Campeol, E. 2000. *In vitro* antimicrobial activity of extracts and isolated constituents of *Geum rivale.* – Phytother. Res., **14**: 561-563.

Petrova, A., Petrović, D., Soljan, D. & Stevanović, V. 2011. *Geum bulgaricum*. - In: IUCN 2013. IUCN Red List of Threatened Species. Version 2013.1

Plischke, A., Choi, Y., Brakefield, P.M., Klinkhamer, P.G.L. & Bruinsma, M. 2012. Metabolomic plasticity in GM and non-GM potato leaves in response to aphid herbivory and virus infection. – J. Agr. Food Chem., **60**: 1488-1493.

Schielmann W, Ammer C, Strack D. 2008. Metabolite profiling of mycorrhizal roots of *Medicago truncatula*. – Phytochemistry, **69**: 112-146.

Schultze, W. & Wollman, C. 1995. Composition of the essential oil from the underground parts of *Geum reptans* L. (*Rosaceae*). – Flav. Fragr. J., **10**: 249-253.

Torras-Claveria, L., Berkov, S., Jáuregui, O., Caujapé, J., Viladomat, F., Codina, C. & Bastida, J. 2010. Metabolite profiling of *Pancratium canariense* extracts: a search for acetylcholinesterase inhibitors. – Phytochem. Analysis, **21**: 80-88.

Tundis, R., Peruzzi, L. & Menichini, F. 2014. Phytochemical and biological studies of *Stachys* species in relation to chemotaxonomy: A review. – Phytochemistry, **102**: 7-39.

Van de Worm, E., Beukelman, C.J., Van de Berq, A.J., Kroes B.H., Labadie, R.P. & Van Dijk, H. 2001. Effects of methoxylation of apocynin and analogs of the inhibition of reactive oxygen species production by stimulated human neutrophils. – Eur. J. Pharmacol., **433**: 225-230.

Wollman, C. & Scbultze, W. 1995. Composition of the root essential oils of several *Geum* species and related members of the subtribus *Geinae* (*Rosaceae*). – Flav. Fragr. J., **10**: 173-178.

Xu, H.-X., Zeng, F.-Q., Wan, M., & Sim, K-Y. 1996. Anti-HIV triterpene acids from *Geum japonicum.* – J. Nat. Prod. (Lloydia), **59**: 643-645.

Xu, H.-X., Ming, D.-Sh., Dong, H. & But, P. 2000. A new anti-HIV triterpene from *Geum japonicum*. – Chem. Pharm. Bull., **48**: 1367-1369.