http://dx.doi.org/10.5935/0103-5053.20160054

J. Braz. Chem. Soc., Vol. 27, No. 10, 1736-1743, 2016.
Printed in Brazil - ©2016 Sociedade Brasileira de Química
0103 - 5053 $6.00+0.00

Article

The HPLC Fingerprint Analysis of Selected *Cirsium* Species with Aid of Chemometrics

Anna M. Hawrył,*,[a] Agata Ziobro,[a] Ryszard S. Swieboda,[a] Mirosław A. Hawrył,[a] Mykhaylo Chernetskyy[b] and Monika Waksmundzka-Hajnos[a]

[a]*Department of Inorganic Chemistry, Medical University of Lublin, 20-093 Lublin, Poland*

[b]*Botanical Garden of Maria Curie-Skłodowska University, 20-081 Lublin, Poland*

Twelve *Cirsium* sp. methanolic extracts were analyzed using high-performance liquid chromatography gradient elution method with run time of 45 min. Four Kinetex (150 × 4.6 mm) chromatographic columns (C18 5 µm, C18 2.6 µm, pentafluorophenyl 5 µm, phenyl-hexyl 5 µm) and mobile phase consisting of methanol/water/formic acid 1% were used. Eight standards (naringin, vanilic acid, chlorogenic acid, caffeic acid, rutin, luteolin, apigenin, *p*-coumaric acid) were analyzed in the same conditions to confirm their presence in all of *Cirsium* methanolic extracts. The obtained chromatograms were compared and the similarity between them was evaluated using the similarity indices (Pearson's correlation coefficient, determination coefficient and congruence coefficient), distance indices (Euclidean, Manhattan and Chebyshev distance) and multi-scale structural similarity (MS-SSIM). Obtained results were confirmed using the principal component analysis (PCA). The attempt of identification of two unknown *Cirsium* species was performed using the similarity, distance indices and PCA analysis.

Keywords: HPLC, fingerprint, *Cirsium*, chemometrics, PCA

Introduction

Cirsium species (Asteraceae) are popular plants growing in the meadows of Europe, North Africa, Siberia, Central Asia and America. In Poland the thistles are very common and widespread. The most popular species are *C. vulgare*, *C. rivulare*, *C. oleraceum*, *C. canum*, *C. eriophorum*, *C. decussatum, C. pannonicum, C. acaule, C. helenoides* and *C. erisithales.* They grow in pastures, fallow, river walleyes and prefer calcareous soils. These plants are famous for their use in traditional and conventional medicine, cosmetology, and some species are used as additive to food because of the nutritional value. The main compounds of *Cirsium* are flavonoids, phenolic acids, sterols, alkaloids, polyacetylenes, acetylenes, triterpenes, sesquiterpene lactones, lignans, hydrocarbons and minerals. The extracts of *Cirsium* exhibit many biological activities, such as antimicrobial,[1,2] anticancer,[3] antioxidant,[4,5] hepatoprotective,[6] antifungal,[7] and antibacterial.[8]

High-performance liquid chromatography (HPLC) fingerprint technique, accepted by World Health Organization (WHO), is widely used for quality control of plant raw material[9] and this method was also used for the study of some *Cirsium* species.[10-14]

The chemometric methods, as the application of mathematical and statistical techniques, can be important instruments to retrieve more information from the chromatographic data and for evaluating of similarity between various herbal species.[15]

In our work, the fingerprint analysis of twelve thistles (ten known and two unknown) were performed using HPLC gradient elution technique. The similarity between the studied *Cirsium* species were evaluated using the similarity indices (Pearson's correlation coefficient, R; determination coefficient, R^2; and congruence coefficient, cosine), distance indices (Euclidean, Manhattan and Chebyshev distances) and multi-scale structural similarity (MS-SSIM). The principal component analysis (PCA) was also necessary to attempt of identification of the two unknown species.

Experimental

HPLC instrumentation and reagents

HPLC analysis was carried out on Hitachi LaChrom Elite System (Tokyo, Japan) with diode array detector L-2455,

*e-mail: anna.hawryl@umlub.pl

thermostat L-2300, pump L-2130 and autosampler L-2200. The chromatographic separation was performed using four Kinetex (Phenomenex, Torrance, CA, USA) chromatographic columns (150 × 4.6 mm): octadecyl carbon chain (C18) bonded silica phase 5 μm, C18 2.6 μm, pentafluorophenyl (PFP) 5 μm and phenyl-hexyl 5 μm, maintained at 30 °C with the run time of 45 min. Detection wavelength was 320 nm with the sample injection volume of 10 μL; the flow rate was 1.0 mL min^{-1}. The gradient elution with gradient concentrations (5-85% v/v) by 45 min with mobile phase consisting of methanol/water/formic acid 1% were used. Methanol Chromasolv for HPLC was purchased from Sigma-Aldrich (St. Louis, MO, USA) and formic acid from Poch (Gliwice, Poland); double-distilled water was used.

Before the analysis, all raw extracts were filtered using filter paper. Ten microliters of 0.1% solutions of standards were applied on the chromatographic column C18 5 μm for the identification of compounds in individual extracts.

Obtained chromatograms of extracts and standards were elaborated using Agilent EZChrom Elite software (Santa Clara, CA, USA).

Extraction procedure

The plant raw material (ten known *Cirsium* species) was obtained from the Botanical Garden of Maria Curie-Skłodowska University (Lublin, Poland) and two unknown plants of the same species were harvested in the meadow in Turka, near Lublin (Poland). The names of the plants are presented in Table 1.

Table 1. Analyzed plants

Sample	Name of plant
1	*Cirsium acaule* Scop.
2	*Cirsium arvense* (L.) Scop.
3	*Cirsium canum* (L.) All.
4	*Cirsium decussatum* Janka
5	*Cirsium eriophorum* (L.) Scop.
6	*Cirsium erisithales* (Jacq.) Scop.
7	*Cirsium helenoides* (L.) Hill
8	*Cirsium pannonicum* (L. fil.) Link
9	*Cirsium rivulare* (Jacq.) All
10	*Cirsium vulgare* (Savi.) Ten.
11	Unknown 1
12	Unknown 2

The identity of species of individual plants was confirmed by the Botanical Garden workers; voucher specimens are placed in the Botanical Garden. The aerial parts of the plants were dried in the shade and wind, at ambient temperature. The mass of raw material was 10 g for samples 1, 2, 3, 7, 9 and 10; 20 g for samples 4, 5 and 8; and 5 g for sample 6. The mass of the two unknown species was 3 g. The aerial parts of dried raw material were ground in a hand mill, then placed in paper case and extracted in a Soxhlet apparatus during 12 h using dichloromethane as solvent and next for another 12 h using methanol as solvent. The obtained extracts were evaporated using rotary vacuum evaporator. Dried extracts were dissolved in methanol and poured into 25 mL graduated flasks. The extracts were stored in the refrigerator.

Preparation of standards

Ten milligrams of eight samples of standards (naringin, vanilic acid, chlorogenic acid, caffeic acid, rutin, luteolin, apigenin and *p*-coumarin acid) were dissolved in 1 mL of methanol to obtain about 0.1% solutions.

Chemometric analysis

Chromatograms of ten known and two unknown species of *Cirsium* were exported to text files (American Standard Code for Information Interchange, ASCII) and then opened using Excel program. The next processing was performed using the data including the retention times and the absorbance values obtained in 320 nm analytical wavelength. The 12 columns (number of studied extracts) and 6750 rows (45 min = 2700 s, frequency of sampling 2.5 Hz) matrix was created and it was saved as .csv format. The obtained file was opened using the program SpecAlign (version 2.4.1),[16] which is often used for alignment process of chromatographic data. Smoothing, denoising and background subtraction are also possible using this program.

At the beginning, the smoothing process was conducted for the obtained chromatograms using the Savitzky-Golay filter. The noise compression was performed using the discrete transformation wavelets Symmlet-8 and next using the soft threshold elimination with the value of threshold parameter equal to 0.5. Then the background subtraction was made. The baseline was designated using the limited moving average method with width of the window equal the twenty percent of chromatogram length.

According to Jiang *et al.*,[17] the recursive alignment by fast Fourier transform (RAFFT) algorithm was selected to chromatograms alignment process. This algorithm is characterized by high efficiency and no effect on peak shape. Aligned chromatogram was divided on the segments and synchronized with target in all segments. The target chromatogram was characterized by highest average correlation coefficient (in this case *C. decussatum*) in

comparison with the others.[18] In all cases, the synchronization was performed with maximum shift equal ten.

After, the similarity and distance indices were calculated and PCA analysis was performed.

Similarity and distance indices

In our work the following similarity and distance indices were used: (*i*) Pearson's correlation coefficient that determines the level of linear dependence between the variables, with values from −1 to 1. A high absolute value of R confirms the strong relationships between the data, and the lack of correlation is when R is equal zero; (*ii*) determination coefficient that determines what percentage of one variable explains the variability of the second one, with R^2 values from 0 to 1. The lack of correlation is observed for zero value and the great similarity is when R^2 is equal 1; (*iii*) congruence coefficient (cosine measure) that is the cosine of the angle between the vectors in n dimensional space. The unit value of congruence coefficient confirms the great similarity between samples; (*iv*) Euclidean distance that is the distance between two points in n dimensional space equal with the length of the segment connecting these points. For similar vectors its value is close to zero; (*v*) Manhattan distance (city block) that is the sum of absolute differences of coordinates pairs of both vectors; (*vi*) Chebyshev distance that is the longest linear segment along one of the directions and it determines the greatest difference of coordinates; (*vii*) MS-SSIM as the plugin for ImageJ program was also used.[19,20] The MS-SSIM was used to calculate the structural multidimensional parameter of similarity for quantitative measure of quality of recognition in optical character recognition (OCR) process. It is based on the picture of the analysis in various scale. Its mathematical definition can be presented as follows:

$$\text{MS-SSIM} = (l_M)^{\alpha_M} \times \Pi_{j=1}^{M}[(c_j)^{\beta_j} \times (s_j)^{\gamma_j}]^{j-1} \quad (1)$$

where M is the greatest coefficient of scale obtained after M-1 iterations. The particular elements of equation such as loss of contrast (c), deformation of lumination (l) and perturbation of the structure (s) are expressed using some indexes determined for all scales separately. The measures of the similarity and the distance were calculated to determine the similarity of analyzed samples and for identification of two unknown thistles.

Results and Discussion

Chromatographic analysis of some extracts of *Cirsium* species were performed by Kozyra and Skalicka-Woźniak[21] and Koryza and Głowniak,[22] and the presence of some flavonoids and phenolic acids in extracts were confirmed.

In our work, the attempt of identification of eight standards (naringin, vanilic acid, chlorogenic acid, caffeic acid, rutin, luteolin, apigenin, *p*-coumarin acid) was performed to confirm their presence in the studied *Cirsium* species (Table 1). The retention times of standards obtained for C18 (5 μm) chromatographic column are presented in Table 2 and the presence of standards in particular studied *Cirsium* extracts are presented in Table 3.

Table 2. Retention times values of standards for C18 (5 μm) column obtained in high-performance liquid chromatography (HPLC)

Standard	Retention time / min
Naryngin	21.53
Vanilic acid	12.05
Chlorogenic acid	12.42
Caffeic acid	12.34
Rutin	22.39
Luteolin	28.78
Apigenin	31.22
p-Coumaric acid	16.41

The retention times of standards in individual chromatograms obtained for C18 (5 μm) and the retention times of substances presented in various *Cirsium* species were compared. The presence of naringin was observed in extracts 1, 2, 4 and 7; vanilic acid in extract 4; chlorogenic acid in samples 2, 3, 5, 8, 9, 10 and in both unknown *Cirsium* species (11 and 12); caffeic acid in samples 1, 7 and 9; rutin was observed in extracts 3, 5, 6, 10 and samples 11 and 12; luteolin in samples 3, 5, 7, 11 and 12; apigenin in extracts 1, 3, 5, 6, 8 and 12; *p*-coumaric acid in samples 1, 2, 3, 6, 7 and 10.

Measures of similarity and distance

Measures of the similarity were calculated for four chromatographic columns (C18 5 μm, C18 2.6 μm, PFP 5 μm, phenyl-hexyl 5 μm) and the summary of results is presented in Table 4. These calculations were performed for chemical comparison of ten analyzed *Cirsium* species and the attempt of identification of two unknown thistles (samples 11 and 12).

The confirmation of identity of unknown *Cirsium* species is ambiguous using these chromatographic and chemometric methods. Our aim was preliminary the estimation of the similarity of studied *Cirsium* species and the attempt of identification of two unknown species.

Table 3. Presence of standards in studied extracts (as detailed in Table 2)

Standard	Number of extract[a]											
	1	2	3	4	5	6	7	8	9	10	11	12
Naryngin	+	+	−	+	−	−	+	−	−	−	−	−
Vanilic acid	−	−	−	+	−	−	−	−	−	−	−	−
Chlorogenic acid	−	+	+	−	+	−	−	+	+	+	+	+
Caffeic acid	+	−	−	−	−	−	+	−	+	−	−	−
Rutin	−	−	+	−	+	+	−	−	−	+	+	+
Luteolin	−	−	+	−	+	−	+	−	−	−	+	+
Apigenin	+	−	+	−	+	+	−	+	−	−	−	+
p-Coumaric acid	+	+	+	−	−	+	+	−	−	+	−	−

[a]In accordance with Table 1.

The similarity between *Cirsium decussatum* Janka (sample 4) and *Cirsium erisithales* (Jacq.) Scop. (sample 6) was noticed in the case of PFP and phenyl-hexyl chromatographic columns using Pearson's correlation coefficient, congruence coefficient (values greater than 0.9) and determination coefficient (values higher than 0.8; Table 4). The similarity between the samples 4 and 5 was confirmed for described above hydrophobic columns, with R and cosine of 0.8857 and 0.8895, respectively, for phenyl-hexyl column; and 0.8298 and 0.8354, respectively, in the case of PFP column. Euclidean, Manhattan and MS-SSIM parameters also confirm the greatest similarity between *Cirsium decussatum* Janka and *Cirsium eriophorum* (L.) Scop.

In the case of both C18 columns, the similarity between samples 4 (*Cirsium decussatum* Janka) or 5 (*Cirsium eriophorum* (L.) Scop.) was observed using all similarity and distance indices. The value of R and cosine are in the range of 0.8652-0.8773; MS-SSIM is equal 0.3322 for C18 2.6 μm and 0.4040 for C18 5 μm. The distance measures (Euclidean, Manhattan and Chebyshev distances) and MS-SSIM also confirm the similarity between samples 4 and 5.

The comparison of two unknown thistles with *Cirsium* species from the Botanical Garden of Maria Curie-Skłodowska University (Lublin, Poland) was performed using the similarity and distance indices. Based on results from Table 4, the similarity between sample 11 and *Cirsium decussatum* Janka (sample 4) was confirmed using the first three similarity parameters (R, R^2 and cosine) for pentafluorophenyl and phenyl-hexyl chromatographic columns. The similarity between samples 11 and 5 (*Cirsium eriophorum* (L.) Scop.) was noticed using R, R^2, cosine and Chebyshev distance (for C18 2.6 μm column) and Chebyshev distance (for PFP and phenyl-hexyl columns).

The similarity of the second unknown *Cirsium* species (sample 12) and *Cirsium canum* (L.) (sample 3) were noticed for four chromatographic columns using R, R^2 and cosine. In the case of PFP and phenyl-hexyl chromatographic columns, the obtained values of Euclidean, Manhattan and Chebyshev distances confirm the similarity between samples 12 and 10 (*Cirsium vulgare* (Savi.) Ten.).

Comparing the two first hydrophobic columns (PFP and phenyl-hexyl), which were used as alternative of octadecyl columns, some differences were observed. It is interesting that the same results (the similarity of 11 and 12 with the other samples) were obtained for these two columns. Sample 11 is similar to *Cirsium decussatum* Janka (sample 4) using the Pearson's correlation coefficient, determination coefficient and congruence coefficient; or to 10, 12, 5 and 2 using appropriately Euclidean, Manhattan, Chebyshev distances and MS-SSIM. The second unknown *Cirsium* species (sample 12) is similar to *Cirsium canum* (L.) All. (3) using R, R^2 and cosine; or to *Cirsium vulgare* (Savi.) Ten. (10) using the three distance indices; or to 11 using the MS-SSIM.

Exemplary chromatograms (for C18 5 μm) obtained for samples 3 (*C. canum*), 12 (unknown species), 4 (*C. decussatum*), 5 (*C. eriophorum*) and 11 (unknown species) are presented in Figure 1.

The summary of all obtained chromatograms for exemplary chromatographic column (C18 5 μm) is presented in Figure 2.

PCA analysis

The preliminary chromatographic data processing, including the smoothing, noise reduction, background subtraction and alignment process, was performed. PCA matrix was consisted of 20 columns and 6751 lines. Obtained results are presented as principal component PC3 and PC2 graphs, indicating the percentage of the variability on the respective axis (Figure 3).

The close proximity of lines corresponding to known and unknown *Cirsium* samples confirms their similarity.

Table 4. Measures of the similarity for four chromatographic columns

No. of extract	R[a]		R^{2b}		Cosine[c]		Distance measure						Other measure	
	Best correlating	Value of R	Best correlating	Value of R^2	Best correlating	Value of cosine	Best correlating	Eucliden	Best correlating	Manhattan	Best correlating	Chebyshev	Best correlating	MS-SSIM
Pentafluorophenyl														
1	6	0.8898	6	0.7917	6	0.8947	4	3775.664	4	72230.02	4	3412350	6	0.322871
2	11	0.7255	11	0.5263	11	0.7291	4	8057.273	6	174756.80	5	5954450	4	0.262962
3	12	0.5471	12	0.2993	12	0.5682	5	6183.705	5	103326.40	8	7505620	5	0.258038
4	6	0.9111	6	0.8301	6	0.9135	5	3593.443	5	61220.03	1	3412350	5	0.402971
5	4	0.8298	4	0.6885	4	0.8354	11	3052.866	4	61220.03	11	2716510	4	0.402971
6	4	0.9111	4	0.8301	4	0.9135	4	5558.137	1	123149.90	4	5220630	1	0.322871
7	11	0.6711	11	0.4504	11	0.6778	5	7266.228	11	129848.40	5	6249313	4	0.179585
8	2	0.5320	2	0.283	2	0.5542	5	6097.550	11	99996.37	2	7104891	4	0.242382
9	6	0.7711	6	0.5946	6	0.7794	1	4392.350	4	84894.33	1	5330360	4	0.251284
10	2	0.5900	2	0.3481	2	0.6162	12	1439.089	12	32305.21	12	1483072	8	0.192871
11	4	0.7790	4	0.6069	4	0.7831	10	2969.790	12	37935.26	5	2716510	2	0.240491
12	3	0.5471	3	0.2993	3	0.5682	10	1439.089	10	32305.21	10	1483072	11	0.154207
Phenyl-hexyl														
1	6	0.8950	6	0.8010	6	0.9960	4	3626.536	4	67400.95	4	3485980	6	0.322871
2	7	0.8020	7	0.6433	7	0.8100	7	6504.800	7	140334.20	5	5718900	4	0.262962
3	12	0.6053	12	0.3664	12	0.6215	5	6460.615	11	95444.94	8	6811910	5	0.258038
4	6	0.9007	6	0.8113	6	0.9037	5	3184.094	5	57579.22	1	3485980	5	0.402971
5	4	0.8857	4	0.7845	4	0.8895	4	3184.094	4	57579.22	11	3425582	4	0.402971
6	4	0.9007	4	0.8113	4	0.9037	4	5755.273	4	130851.40	4	5272340	1	0.322871
7	2	0.8020	2	0.6433	2	0.8100	2	6504.800	11	131598.50	2	6088572	4	0.179585
8	2	0.6243	2	0.3897	2	0.6410	5	5312.688	11	89252.25	2	6027250	4	0.242382
9	6	0.7876	6	0.6202	6	0.7955	1	4154.353	4	87064.68	1	4874350	4	0.251284
10	2	0.7402	2	0.5479	2	0.7555	12	1599.891	12	33997.20	12	1718599	8	0.192871
11	4	0.7858	4	0.6175	4	0.7894	10	2958.897	12	38305.16	5	3425582	2	0.240491
12	3	0.6053	3	0.3664	3	0.6215	10	1599.891	10	33997.20	10	1718599	11	0.154207
C18 2.6 µm														
1	4	0.8338	4	0.6952	4	0.8398	4	326.512	4	65862.62	4	3053300	6	0.282646
2	11	0.7300	11	0.5329	5	0.7416	4	6883.384	4	156458.20	4	5134449	1	0.232315
3	7	0.5595	7	0.3130	7	0.5778	5	5964.378	11	101460.80	7	5678681	9	0.225423
4	5	0.8728	5	0.7618	5	0.8773	5	3013.701	5	59202.68	5	2906770	5	0.332169
5	4	0.8728	4	0.7618	4	0.8773	11	2592.399	11	58176.19	11	2275616	4	0.332169
6	1	0.7266	1	0.5279	1	0.7408	2	13128.100	2	377784.00	4	7511734	1	0.282646
7	2	0.6983	2	0.4877	2	0.7114	5	7220.450	11	133710.00	5	5210367	1	0.191713
8	10	0.7113	10	0.5060	10	0.7223	5	5584.425	11	86677.43	4	5871000	4	0.228071
9	4	0.7022	4	0.4931	4	0.7122	1	4344.420	4	85596.44	1	5194280	3	0.225423
10	8	0.7113	8	0.5060	8	0.7223	12	1422.681	12	32719.69	12	1346388	4	0.175540
11	5	0.8221	5	0.6758	5	0.8237	10	2403.386	12	30635.25	5	2275616	2	0.196416
12	3	0.4629	3	0.2143	3	0.4918	10	1422.681	11	30635.25	10	1346388	8	0.127821
C18 5 µm														
1	6	0.9176	6	0.8420	6	0.9223	6	5039.912	6	113231.80	6	5455930	2	0.236204
2	5	0.7353	5	0.5407	5	0.7473	7	7124.137	7	147520.20	5	5504750	1	0.236204
3	12	0.7515	12	0.5647	12	0.7625	5	5591.485	5	96856.37	7	6997140	9	0.201067
4	5	0.8652	5	0.7486	5	0.8696	5	2950.510	5	53762.34	5	3029820	5	0.404037
5	4	0.8652	4	0.7486	4	0.8696	4	2950.510	4	53762.34	4	3029820	4	0.404037
6	1	0.9176	1	0.8420	1	0.9223	1	5039.912	1	113231.80	1	5455930	4	0.242998
7	2	0.7328	2	0.5370	2	0.7420	5	6855.439	4	128350.20	5	6188920	9	0.167096
8	2	0.4830	2	0.2333	2	0.5051	5	6252.708	12	110410.00	3	7095652	4	0.198707
9	6	0.7155	6	0.5119	6	0.7243	4	4830.872	4	85101.52	4	5767650	1	0.218126
10	2	0.4061	2	0.1649	2	0.4396	5	7222.264	12	159578.80	5	5823470	9	0.054843
11	6	0.7003	6	0.4905	6	0.7041	4	5374.241	12	65437.83	4	6355910	2	0.196578
12	3	0.7515	3	0.5647	3	0.7625	5	4636.267	11	65437.83	5	5939895	9	0.091311

[a]Pearson's correlation coefficient; [b]determination coefficient; [c]congruence coefficient. MS-SSIM: Multi-scale structural similarity.

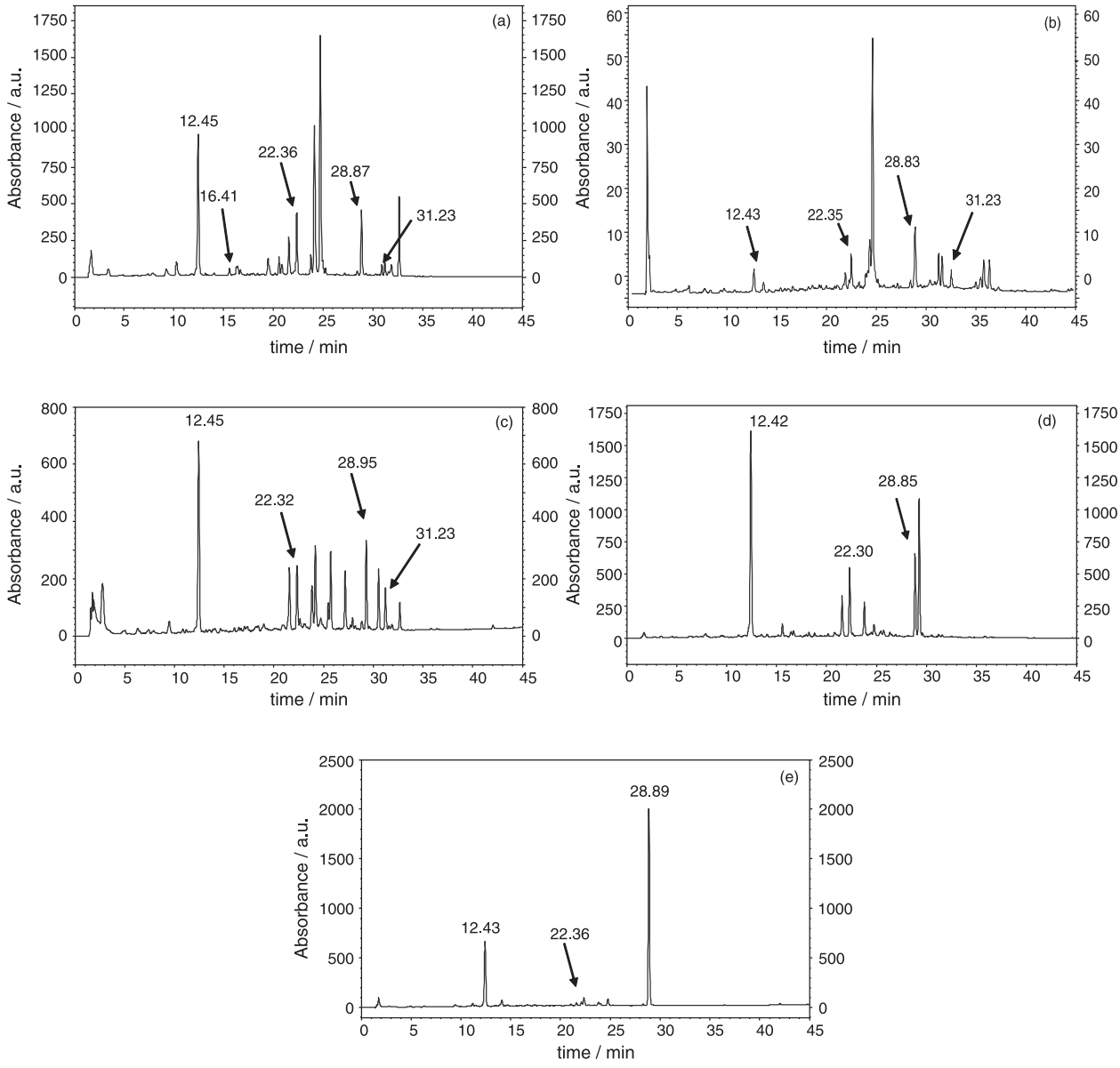

Figure 1. High-performance liquid chromatograph (HPLC) chromatograms for C18 5 μm column of (a) *Cirsium canum* extract; (b) unknown *Cirsium* species (sample 12); (c) *Cirsium eriophanum* extract; (d) *Cirsium decussatum* extract; and (e) unknown *Cirsium* species (sample 11). Retention times values of standards in accordance with Table 2.

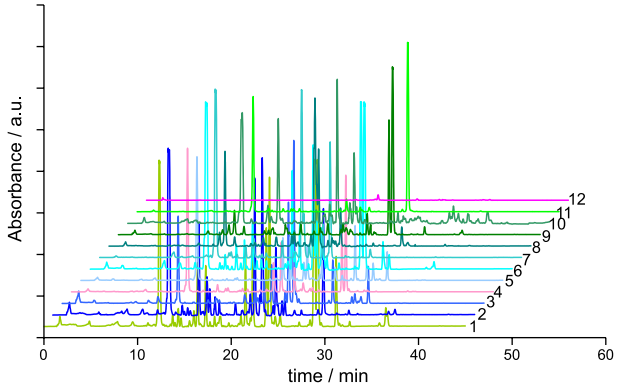

Figure 2. Summary of chromatograms for C18 (5 μm) column. Numbers of extracts as in Table 1.

The similarity of *Cirsium decussatum* Janka (sample 4) and *Cirsium erisithales* (Jacq.) Scop. (sample 6) was also confirmed by PCA analysis. In the PC2 *vs.* PC3 charts for PFP and C18 (5 μm) columns (Figures 3a and 3d) lines corresponding with samples 4 and 6 are close to each other. For other chromatographic columns (Figures 3b and 3c) these lines are also near to each other, but not so close.

The similarity of sample 11 and *Cirsium arvense* (L.) Scop. (2) or *Cirsium helenoides* (L.) Hill (7) was observed (lines corresponding to particular samples are close to each other) for PFP column (Figure 3a). In case of phenylhexyl column (Figure 3b), the similarity of sample 11 with samples 2, 5, 7 and 10 (the nearest line) was observed.

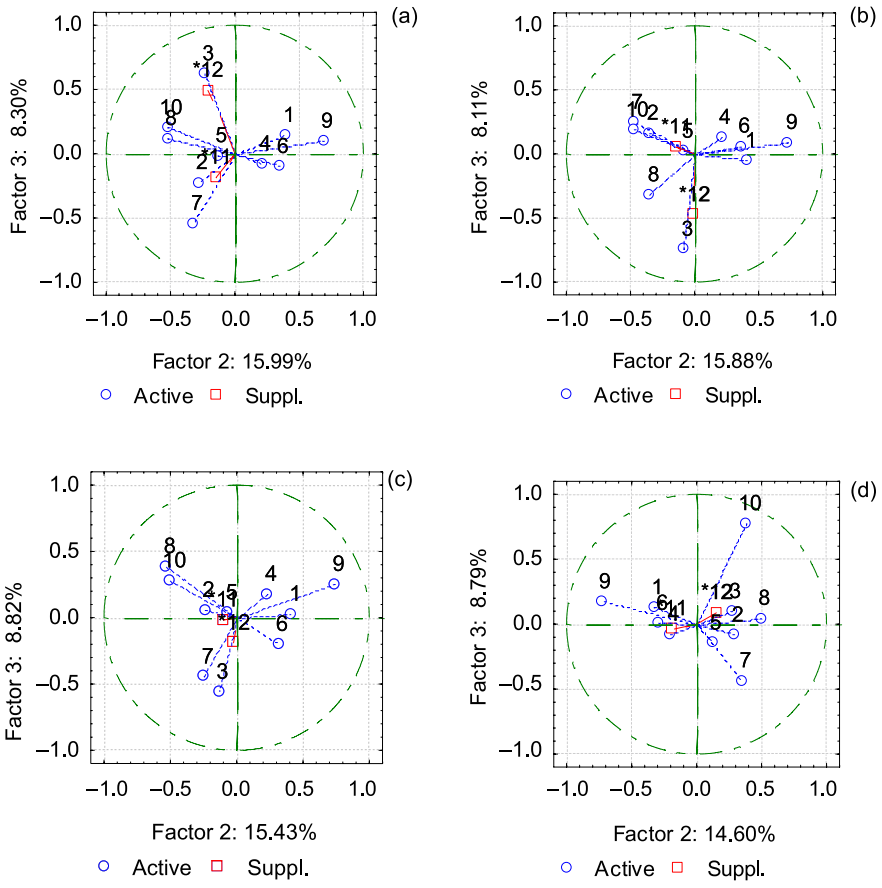

Figure 3. Principal component analysis (PCA) graphs for data matrix of (a) pentafluorophenyl; (b) phenyl-hexyl; (c) C18 2.6 µm; and (d) C18 5 µm. The circles (active) are the known *Cirsium* species (samples 1-10), and the squares (suppl.) are the unknown *Cirsium* species (samples 11 and 12).

For C18 (2.6 µm) chromatographic columns (Figure 3c), the line corresponding with the first unknown *Cirsium* species (11) is close to 2 (*Cirsium arvense* (L.) Scop.) and 5 (*C. eriophorum*), but for C18 (5 µm) 11 is located close to 4 (*Cirsium decussatum* Janka) and 6 (*Cirsium erisithales* (Jacq.) Scop.).

The second unknown *Cirsium* sp. (sample 12) was also compared to the other known *Cirsium* sp. For all chromatographic columns (Figures 3a-d), the line corresponding to sample 12 is located near sample 3 (*Cirsium canum*).

In conclusion, the similarity between samples 12 and 3 (*Cirsium canum*) was confirmed using the similarity indices (R, R^2 and cosine) and PCA for all used chromatographic systems. Moreover, the similarity between unknown sample 12 and *Cirsium vulgare* (Savi.) Ten. (sample 10) was confirmed using the distance parameters (Euclidean, Manhattan and Chebyshev distances) for PFP and phenyl-hexyl columns. The similarity between the second unknown *Cirsium* species (sample 11) and sample 2 (*Cirsium arvense* (L.) Scop.) was confirmed using MS-SSIM parameter for all chromatographic systems; and for PFP, phenyl-hexyl and C18 (2.6 µm) columns using PCA method. The similarity of sample 11 and 4 (PFP and phenyl-hexyl), 5 (C18 2.6 µm) and 6 (C18 5 µm) was confirmed using the similarity indices (R, R^2 and cosine); whereas the distance parameters confirm the similarity between 11 and 10, 12, 5 and 2 (for PFP, phenyl-hexyl and C18 2.6 µm) and 4, 12 and 2 (for C18 5 µm). The PCA analysis confirms the similarity between 11 and 2 for the first three chromatographic columns.

Conclusions

The chromatographic fingerprint constructions of twelve *Cirsium* species were prepared using HPLC and chemometric methods. The attempt of identification of standards (naringin, vanilic acid, chlorogenic acid, caffeic acid, rutin, luteolin, apigenin, *p*-coumaric acid) was performed based on the retention time values of particular standards. The presence of some standards in all *Cirsium* methanolic extracts was confirmed.

The similarity between various *Cirsium* species was evaluated using the similarity and distance indices. The similarity of unknown *Cirsium* species (12) and *Cirsium canum* was confirmed using the similarity indices

(R, R^2 and cosine) whereas the distance parameters (Euclidean, Manhattan and Chebyshev distances) confirm the similarity between unknown sample 12 and *Cirsium vulgare* (Savi.) Ten. for PFP and phenyl-hexyl columns. PCA analysis conforms the similarity between 12 and *Cirsium canum* for all used chromatographic systems.

The MS-SSIM parameter confirms the similarity between the second unknown *Cirsium* species (sample 11) and *Cirsium arvense* (L.) Scop. for all chromatographic systems. The PCA analysis also confirms it in case of PFP, phenyl-hexyl and C18 2.6 μm columns. Moreover, the similarity indices (R, R^2 and cosine) confirm the similarity of sample 11 and *Cirsium decussatum* Janka (PFP and phenyl-hexyl), *Cirsium eriophorum* (L.) Scop. (C18 2.6 μm) and *Cirsium erisithales* (Jacq.) (C18 5 μm); whereas the distance parameters (Euclidean, Manhattan and Chebyshev) confirm the similarity between 11 and *Cirsium vulgare* (Savi.) Ten., the second unknown species, *Cirsium eriophorum* (L.) and *Cirsium arvense* (L.) Scop. (for PFP, phenyl-hexyl and C18 2.6 μm) and *Cirsium decussatum* Janka, the second unknown species and *Cirsium arvense* (L.) Scop. (for C18 5 μm). In the case of the PFP, phenyl-hexyl and C18 2.6 μm, the PCA analysis confirms the similarity between 11 and *Cirsium arvense* (L.) Scop.

Acknowledgments

The authors would like to thank PhD Grażyna Szymczak (director of the Botanical Garden of Maria Skłodowska-Curie University, Lublin, Poland) for the acquisition of raw material.

References

1. Nazaruk, J.; Jakoniuk, P.; *J. Ethnopharmacol.* **2005**, *102*, 208.
2. Karasakal, A.; Demirci, A. S.; *Marmara Pharm. J.* **2015**, *19*, 43.
3. Liu, S.; Zhang, J.; Li, D.; Liu, W.; Luo, X.; Zhang, R.; Li, L.; Zhao, J.; *Nat. Prod. Res.* **2007**, *21*, 915.
4. Nazaruk, J.; Czechowska, S. K.; Markiewicz, R.; Borawska, M. H.; *Nat. Prod. Res.* **2008**, *22*, 1583.
5. Nazaruk, J.; Wajs-Bonikowska, A.; Bonikowski, R.; *Chem. Nat. Compd.* **2012**, *48*, 8.
6. Kuo-Lung, K.; Chu-Tsai, T.; Wei-Min, C.; Mei-Lin, S.; Chia-Tien, W.; Hui-Fen, L.; *Am. J. Chin. Med.* **2008**, *36*, 355.
7. Yoon, G. J.; Choi, Y. H.; Choi, K. S.; Jang, B.; Cha, J.-C.; Kim, M.-Y.; *Ind. Crops Prod*. **2011**, *34*, 882.
8. Kozyra, M.; Biernasiuk, A.; Malm, A.; Chowaniec, M.; *Nat. Prod. Res*. **2015**, *29*, 2059.
9. Tistaert, C.; Dejaegher, B.; Vander Heyden, Y.; *Anal. Chim. Acta* **2011**, *690*, 148.
10. Lu, Y.; Song, W.; Liang, X.; Wei, D.; Zhou, X.; *Chromatographia* **2009**, *70*, 125.
11. Ni, X.; Xu, Q.; *J. Pharm. Anal.* **2007**, *6*, 929.
12. Ganzera, M.; Pöcher, A.; Stuppner, H.; *Phytochem. Anal.* **2005**, *16*, 205.
13. Jeong, D. M.; Jung, H. A.; Choi, J. S.; *Arch. Pharmacal Res.* **2008**, *31*, 28.
14. Thao, N. T.; Cuong, T. D.; Hung, T. M.; Lee, J. H.; Na, M.; Son, J. K.; Jung, H. J.; Fang, Z.; Woo, M. H.; Choi, J. S.; Min, B. S.; *Arch. Pharmacal Res.* **2011**, *34*, 455.
15. Bansal, A.; Chhabra, V.; Rawal, R. K.; Sharma, S.; *J. Pharm. Anal.* **2014**, *4*, 223.
16. Wong, J. W. H.; Cagney, G.; Cartwright, H. M.; *Bioinformatics* **2005**, *21*, 2088.
17. Jiang, W.; Zhang, Z.; Yun, Y.; Zhan, D.; Zheng, Y.; Liang, Y.; Yang, Z.; Yu, L.; *Chromatographia* **2013**, *76*, 1067.
18. Daszykowski, M.; Walczak, B.; *J. Chromatogr. A* **2007**, *1176*, 1.
19. http://rsb.info.nih.gov/ij/ accessed in February 2016.
20. Rouse, D.; Hemami, S.; Analyzing the Role of Visual Structure in the Recognition of Natural Image Content with Multi-Scale SSIM; *Proceedings of SPIE*, 6806, Human Vision and ElectronicImaging XIII, San Jose, CA **2008**.
21. Kozyra, M.; Skalicka-Woźniak, K.; *Acta Pol. Pharm.* **2014**, *71*, 877.
22. Kozyra, M.; Głowniak, K.; *Acta Soc. Bot. Pol.* **2013**, *82*, 325.

Submitted: October 14, 2015
Published online: February 17, 2016