SUPPLEMENTARY MATERIAL

Secoiridoids and other chemotaxonomically relevant compounds in *Pedicularis*: phytochemical analysis and comparison of *Pedicularis rostratocapitata* Crantz and *Pedicularis verticillata* L. from Dolomites

Alessandro Venditti^{*a,b}, Claudio Frezza^b, Fabio Sciubba^a, Sebastiano Foddai^b, Mauro Serafini^b,
Marcello Nicoletti^b, Armandodoriano Bianco^a

^{a)} Dipartimento di Chimica, Sapienza Università di Roma, P.le A. Moro 5, 00185 Roma, Italy,

^{b)} Dipartimento di Biologia Ambientale, Sapienza Università di Roma,

*E-mail: alessandro.venditti@uniroma1.it

* Corresponding author. Tel. +39 0649913229; fax: +39 0649913841;

E-mail address: alessandro.venditti@uniroma1.it (A. Venditti).

Abstract.

We compared the respective metabolite patterns of two *Pedicularis* species from Dolomites. Seven phenylethanoid glycosides, i.e. verbascoside (**1**), echinacoside (**2**), angoroside A (**3**), cistantubuloside B1 (**4**), wiedemannioside C (**5**), campneoside II (**11**) and cistantubuloside C1 (**12**), together with several iridoid glucosides as aucubin (**6**), euphroside (**7**), monomelittoside (**8**), mussaenosidic acid (**9**) and 8-epiloganic acid (**13**) were identified. *P. verticillata* showed also the presence of greatly unexpected secoiridoids, ligustroside (**14**) and excelside B (**15**), very rare compounds in Lamiales. Both PhGs and iridoids are considered of taxonomical relevance in the Asteridae and their occurrence in *Pedicularis* was discussed. In particular, the exclusive presence of several compounds such as 8-epiloganic acid (**13**), campneoside II (**11**), cistantubuloside C1 (**12**), ligustroside (**14**) and excelside B (**15**) in *P. rostratocapitata*, and angoroside A (**3**), cistantubuloside B1 (**4**) and wiedemannioside C (**5**) in *P. verticillata* could be considered specific markers for the two botanical entities.

Keywords: *Pedicularis rostratocapitata*, *Pedicularis verticillata*, Orobanchaceae, iridoid glucosides, phenylethanoid glycosides, secoiridoids, chemotaxonomy.

3 - Experimental.

3.1 - General.

NMR spectra were recorded on Varian (now Agilent Technologies, Santa Clara, CA, USA) Mercury 300 MHz and/or on Bruker Avance III 400 MHz instrument (Bruker, Billerica, MA, USA) using CD₃OD or D₂O as deuterated solvents; the chemical shift was expressed in ppm (the signal of HDO (s) at 4.78 ppm was used as reference for spectra in D₂O; and the internal solvent signal (m5) at 3.31 ppm for spectra recorded in CD₃OD). MS spectra were performed on a Q-TOF MICRO spectrometer (Micromass, now Waters, Manchester, UK) equipped with an ESI source, that operated in the negative and/or positive ion mode. The flow rate of sample infusion was 10 µL/min. with 100 acquisitions per spectrum. Data were analysed by using the MassLynx software developed by Waters. Solvents of RPE grade were purchased from Sigma Aldrich or Carlo Erba Reagenti, silica gel 60 (70-230 mesh ASTM) from Fluka.

3.2 - Plant Materials.

The examined samples were collected between 2000 and 2500 m a.s.l. on Monte Paterno (Paternkofel) in the Dolomites Massif, Italy. The two species were sampled at the beginning of August when in flowering. *P. verticillata* was collected from an alpine pasture on calcareous/siliceous substrate, *P. rostratocapitata* was collected on a slope among rocks and debris on calcareous substrate. To conciliate the need to have a representative sample of the two species and the impossibility to collect a large amount of plant materials from plants growing in protected areas, we have taken a single branch from each plant, sampling five/six different plants for each species. The botanical identification of the species was performed by one of us (A.V.) using available literature (Pignatti, 1982; Aeschimann et al., 2004; Conti et al, 2005) . The specimens of the studied plants were registered at “Museo Erbario RO, Dipartimento di Biologia Vegetale, Sapienza Università di Roma” in the “Anzalone Collection” under the following accession numbers: 34528 for *P. verticillata* and 34561 for *P. rostratocapitata*.

3.3 - Extraction and isolation of polar compounds.

The plant materials (aerial parts) were dried in a desiccator at room temperature and then were extracted with ethanol 96% four time (200 mL each extraction). The extracts were collected and the ethanol was eliminated at reduced pressure until a water suspension was obtained. The suspension was frozen and then lyophilized, to preserve temperature-sensitive compounds, obtaining the dried crude extract which was subjected to repeated chromatographic fractionation on silica gel column, using different eluting mixtures, as *n*-butanol saturated with water, *n*-butanol/methanol/water

(63:10:27) and mixtures of chloroform/methanol at various percentages, and varying the polarity during the chromatographic run from 95:5 to 6:4.

Were extracted 17.7 g of *P. verticillata* and 12.5 g of *P. rostratocapitata* dried materials obtaining respectively 2.10 and 0.91 g of crude extracts, respectively. The chromatographic separation protocol was the same used in our recent study on *Pedicularis kernerii* (Venditti et al, 2015b in press).

3.4 - Bidimensional NMR experiments

Bidimensional spectra were performed on a Bruker Avance III 400 MHz instrument, operating at 9.4 T at 298° K. TOCSY experiments were acquired with a spectral width of 15 ppm in both dimensions, a mixing time length of 110 ms, recycle delay of 2 s and a data matrix of 4K x 256 points. NOESY experiments were acquired with a spectral width of 15 ppm in both dimensions, a mixing time length of 700 ms, recycle delay of 2 s and a data matrix of 4K x 256 points. HSQC experiments were acquired with a spectral width of 15 and 250 ppm for the proton and carbon respectively, an average $^1J_{C-H}$ of 145 Hz, recycle delay of 2 s and a data matrix of 4K x 256 points. HMBC experiments were acquired with a spectral width of 15 and 250 ppm for the proton and carbon respectively, long range coupling constant of $^nJ_{C-H}$ of 8 Hz, recycle delay of 2 s and a data matrix of 4K x 256 points.

3.5 - NMR data of isolated compounds

verbascoside (**1**) : 1H NMR (300 MHz, CD_3OD) δ : 7.60 (1H, d, J = 15.9 Hz, H- β_{Caff}), 7.06 (1H, d, J = 1.7 Hz, H-2 $_{Caff}$), 6.96 (1H, dd, J = 8.1, 1.7 Hz, H-6 $_{Caff}$), 6.78 (1H, d, J = 8.2 Hz, H-5 $_{Caff}$), 6.72 – 6.65 (2H, overlapped, H-2 $_{Tyr}$ / H-5 $_{Tyr}$), 6.57 (1H, dd, J = 8.0, 1.8 Hz, H-6 $_{Tyr}$), 6.28 (1H, d, J = 15.9 Hz, H- α_{Caff}), 5.19 (1H, *br s*, H-1 $_{Rha}$), 4.93 (1H, partially overlapped with HDO signal, H-4 $_{Glc}$), 4.38 (1H, d, J = 7.9 Hz, H-1 $_{Glc}$), 2.79 (2H, t, J = 5.5 Hz, H- β_{Tyr}), 1.09 (3H, d, J = 6.1 Hz, H-6 $_{Rha}$).

ESI-MS: m/z 646.96 $[M+Na]^+$

echinacoside (**2**): Data reported in Supplementary Table 1.

angoroside A (**3**): 1H NMR (300 MHz, CD_3OD) δ : 7.60 (1H, d, J = 16.0 Hz, H- β_{Caff}), 7.06 (1H, d, J = 1.8 Hz, H-2 $_{Caff}$), 6.96 (1H, dd, J = 7.9, 1.8 Hz, H-6 $_{Caff}$), 6.78 (1H, d, J = 8.0 Hz, H-5 $_{Caff}$), 6.73 – 6.64 (2H, m, H-2 $_{Tyr}$ /H-6 $_{Tyr}$), 6.58 (1H, d, J = 8.2 Hz, H-5 $_{Tyr}$), 6.27 (1H, d, J = 16.1 Hz, H- α_{Caff}), 5.19 (1H, d, J = 1.3 Hz, H-1 $_{Rha}$), 4.38 (1H, d, J = 7.8 Hz, H-1 $_{Glc}$), 4.30 (1H, d, J = 6.8 Hz, H-1 $_{ara}$), 2.80 (2H, t, J = 7.2 Hz, H- β_{Tyr}), 1.09 (3H, d, J = 6.0 Hz, H-6 $_{Rha}$).

ESI-MS: m/z 778.93 $[M+Na]^+$

cistantubuloside B₁ (**4**): 1H NMR (400 MHz, D₂O) δ : 7.88 (2H, d, J = 8.1 Hz, H-2_{Cou}/H-6_{Cou}), 7.78 (1H, d, J = 16.0 Hz, H- β _{Cou}), 6.98 (2H, d, J = 8.1 Hz, H-3_{Cou}/H-5_{Cou}), 6.90 (1H, d, J = 8.0 Hz, H-5_{Tyr}), 6.89 (1H, d, J = 2.0 Hz, H-2_{Tyr}), 6.79 (1H, dd, J = 8.1, 1.9 Hz, H-6_{Tyr}), 6.49 (1H, d, J = 15.9 Hz, H- α _{Cou}), 5.15 (1H, d, J = 1.6 Hz, H-1_{Rha}), 5.06 (1H, br t, J = 6.9 Hz, H-4_{Glc}), 4.54 (1H, d, J = 8.1 Hz, H-1_{innerGlc}), 4.39 (1H, d, J = 7.8 Hz, H-1_{outerGlc}), 2.86 (2H, br t, J = 6.8 Hz, H- β _{Tyr}), 1.08 (3H, d, J = 6.2 Hz, H-6_{Rha}).

ESI-MS: m/z 792.93 $[M+Na]^+$

wiedemannioside C (**5**): 1H NMR (300 MHz, D₂O) δ : 7.67 (1H, d, J = 15.6 Hz, H- β _{Fer}), 7.17 (1H, br s, H-2_{Fer}), 7.10 (1H, br d, J = 8.5 Hz, H-6_{Fer}), 6.90 (1H, d, J = 8.5 Hz, H-5_{Fer}), 6.82 (2H, m, H-2_{Tyr}/H-5_{Tyr}), 6.72 (2H, br d, J = 8.0 Hz, H-6_{Tyr}), 6.37 (1H, d, J = 15.6 Hz, H- α _{Fer}), 5.08 (1H, br s, H-1_{Rha}), 5.05 (1H, t, J = 9.2 Hz, H-4_{innerGlc}), 4.48 (1H, d, J = 7.9 Hz, H-1_{outerGlc}), 4.32 (1H, d, J = 8.0 Hz, H-1_{innerGlc}), 3.79 (3H, s, $\underline{CH_3O^-}$), 2.80 (2H, br t, J = 7.1 Hz, H- β _{Tyr}), 1.01 (3H, d, J = 6.1 Hz, H-6_{Rha}).

ESI-MS: m/z 822.91 $[M+Na]^+$

aucubin (**6**): 1H NMR (400 MHz, D₂O) δ : 6.31 (1H, dd, J = 6.1, 1.7 Hz, H-3), 5.86 (1H, d, J = 1.6 Hz, H-7), 5.26 (1H, d, J = 5.2 Hz, H-1), 5.12 (1H, dd, J = 6.2, 3.6 Hz, H-4), 4.83 (d, J = 8.1 Hz, H-1), 4.54 (1H, m, H-6), 4.31 (2H, dd, J = 17.6, 14.9 Hz, H-10), 3.15 (1H, t, J = 6.1 Hz, H-9), 2.82 – 2.75 (1H, m, H-5).

ESI-MS: m/z 369.00 $[M+Na]^+$

euphroside (**7**): 1H NMR (400 MHz, D₂O) δ : 9.22 (1H, s, H-11), 7.53 (1H, s, H-3), 5.99 (1H, s, H-1), 4.86 (1H, d, J = 8.0 Hz, H-1'), 2.54 (1H, br s, H-9), 2.32 – 1.65 (4H, m, H-6/H-7), 1.22 (3H, s, H-10).

^{13}C NMR (100 MHz, D₂O) δ : 193.7 (C-11), 163.5 (C-3), 126.9 (C-4), 98.6 (C-1'), 95.9 (C-1), 78.7 (C-8), 78.2 (C-3'), 76.3 (C-5'), 74.1 (C-2'), 71.4 (C-5), 63.9 (C-9), 63.4 (C-6'), 38.5 (C-7), 35.9 (C-6), 22.2 (C-10).

ESI-MS: m/z 399.01 $[M+Na]^+$; m/z 375.09 $[M-H]^-$

monomelittoside (**8**): 1H NMR (300 MHz, D₂O) δ : 6.38 (1H, d, J = 6.5 Hz, H-3), 5.89 (1H, br s, H-7), 5.73 (1H, br s, 1H), 5.07 (1H, d, J = 6.5 Hz, H-4), 4.25 (2H, dd, J = 15.0, 8.2 Hz, H-10), 3.25 (1H, m, H-9).

ESI-MS: m/z 384.98 $[M+Na]^+$

mussaenosidic acid (**9**): 1H NMR (400 MHz, D_2O) δ : 7.42 (1H s, H-3), 5.60 (1H, d, $J = 2.7$ Hz, H-1), 2.39 (1H, dd, $J = 9.7, 2.7$ Hz, H-9), 2.26 – 1.60 (4H, m, H-6/H-7), 1.34 (3H, s, H-10).

^{13}C NMR (100 MHz, D_2O) δ : 169.9 (C-11), 150.7 (C-3), 111.6 (C-4), 98.3 (C-1'), 93.0 (C-1), 80.7 (C-8), 78.7 (C-5'), 76.2 (C-3'), 74.4 (C-2'), 70.8 (C-4'), 60.69 (C-6'), 51.8 (C-9), 40.8 (C-7), 30.16 (C-5), 29.7 (C-6), 22.2 (C-10).

ESI-MS: m/z 375.14 $[M-H]^-$; m/z 398.99 $[M+Na]^+$; m/z 414.97 $[M+K]^+$.

mannitol (**10**): ^{13}C NMR (100 MHz, D_2O) δ : 70.83 (C-2/C-5), 69.27 (C-3/C-4), 63.21 (C-1/C-6).

ESI-MS: m/z 205.00 $[M+Na]^+$; m/z 181.05 $[M-H]^-$

campneoside II (**11**): 1H NMR (300 MHz, CD_3OD) δ 7.52 (1H, d, $J = 15.9$ Hz, H- β_{Caff}), 6.98 (1H, d, $J = 2.0$ Hz, H-2 $_{Caff}$), 6.88 (1H, dd, $J = 8.1, 2.0$ Hz, H-6 $_{Caff}$), 6.70 (1H, d, $J = 8.2$ Hz, H-5 $_{Caff}$), 6.63 – 6.56 (2H, m, H-2 $_{Tyr}$ /H-5 $_{Tyr}$), 6.49 (1H, dd, $J = 8.0, 1.8$ Hz, H-6 $_{Tyr}$), 6.20 (1H, d, $J = 15.9$ Hz, H- α_{Caff}), 5.30 (1H, d, $J = 1.9$ Hz, H-1 $_{Rha}$), 4.75 (1H, overlapped with HDO signal, H-4 $_{Glc}$), 4.40 – 4.35 (1H, m, H- α_{Tyr}), 4.30 (1H, d, $J = 7.9$ Hz, H-1 $_{Glc}$), 3.90 – 3.78 (2H, m, H- α_{Tyr}), 1.02 (3H, d, $J = 6.1$ Hz, H-6 $_{Rha}$).

ESI-MS: m/z 662.93 $[M+Na]^+$; m/z 1303.65 $[2M+Na]^+$

Cistantubuloside C₁ (**12**): 1H NMR (400 MHz, D_2O) δ : 7.73 (1H, d, $J = 15.9$ Hz, H- β_{Caff}), 7.24 (1H, d, $J = 1.9$ Hz, H-2 $_{Caff}$), 7.17 (1H, dd, $J = 8.2, 1.9$ Hz, H-6 $_{Caff}$), 6.97 (1H, d, $J = 8.2$ Hz, H-5 $_{Caff}$), 6.90 (1H, d, $J = 8.2$ Hz, H-5 $_{Tyr}$), 6.89 (1H, d, $J = 2.1$ Hz, H-6 $_{Tyr}$), 6.79 (1H, dd, $J = 8.2, 2.1$ Hz, H-6 $_{Tyr}$), 6.43 (1H, d, $J = 16.0$ Hz, H- α_{Caff}), 5.15 (1H, d, $J = 1.6$ Hz, H-1 $_{Rha}$), 5.06 (1H, t, $J = 9.7$ Hz, H-4 $_{innerGlc}$), 4.54 (1H, d, $J = 8.1$ Hz, H-1 $_{innerGlc}$), 4.39 (1H, d, $J = 7.9$ Hz, H-1 $_{outerGlc}$), 1.08 (3H, d, $J = 6.2$ Hz, H-6 $_{Rha}$).

ESI-MS: m/z 825.12 $[M+Na]^+$

8-epiloganic acid (**13**): 1H NMR (300 MHz, D_2O) δ : 7.37 (1H, s, H-3), 5.59 (1H, br s, H-1), 4.68 (1H, d, $J = 7.7$ Hz, H-1'), 3.14 – 2.99 (1H, m, H-5), 2.84 – 2.72 (1H, m, H-9), 2.22 (1H, m, H-8), 2.02-1.88 (2H, m, H-6), 1.04 (3H, d, $J = 6.7$ Hz, H-10).

^{13}C NMR (75 MHz, D_2O) δ : 170.1 (C-11), 142.9 (C-3), 115.0 (C-4), 98.0 (C-1'), 97.2 (C-1), 76.0 (C-3'), 75.9 (C-5'), 75.6 (C-7), 73.9 (C-2'), 70.8 (C-4'), 62.8 (C-6'), 43.2 (C-9), 41.6 (C-6), 41.3 (C-8), 36.6 (C-5), 23.2 (C-10).

ESI-MS: m/z 375.11 $[M-H]^-$; m/z 399.14 $[M+Na]^+$; 776.41 $[2M+Na]^+$.

Ligustroside (**14**): 1H NMR (400 MHz, D₂O) δ : 7.61 (1H, s, H-3), 7.20 (2H, d, J = 8.5 Hz, H-4''/H-8''), 6.85 (2H, d, J = 8.5 Hz, H-5''/H-7''), 6.13 (1H, q, J = 7.5 Hz, H-8), 5.93 (1H, br s, H-1), 4.95 (1H, d, J = 8.0 Hz, Glc-H-1'), 3.72 (3H s, COOCH₃), 2.83 (1H, dd, J = 13.6, 4.1 Hz, H-6a), 2.56 (1H, dd, J = 13.6, 8.5 Hz, H-6b), 1.74 (3H, d, J = 7.0 Hz, H-10).

ESI-MS: m/z 523.43 $[M-H]^-$; m/z 547.45 $[M+Na]^+$;

Excelside B (**15**): 1H NMR (400 MHz, D₂O) δ : 7.69 (1H, s, H-3), 7.38 (2H, d, J = 8.4 Hz, H-4''/H-8''), 7.04 (2H, d, J = 8.4 Hz, H-5''/H-7''), 6.24 (1H, q, J = 7.2 Hz, 1H), 6.07 (1H, br s, H-1), 4.98 (1H, d, J = 7.8 Hz, H-1'), 4.52 (1H, d, J = 8.0 Hz, H-1''), 3.93 (1H, dd, J = 12.3, 1.7 Hz, H-6'a), 3.80 (3H, s, COOCH₃), 3.76 (1H, dd, J = 12.5, 6.9 Hz, H-6'b), 2.80 (1H, dd, J = 13.9, 4.9 Hz, H-6a), 2.56 (1H, dd, J = 13.9, 8.4 Hz, H-6b), 1.78 (3H, d, J = 7.0 Hz, H-10).

ESI-MS: m/z 685.57 $[M-H]^-$; m/z 709.20 $[M+Na]^+$;

Table S1: NMR data of echinacoside (**2**) in D₂O.

Moiety	δ 1H	Multiplicity (Hz)	δ ^{13}C	Intermoiety HMBC Correlations	Intermoiety NOESY Correlations
Aglycone					
1	\	\	132.8		
2	6.80	d (1.9)	118.0		
3	\	\	148.8		
4	\	\	143.6		
5	6.70	dd (5.5, 1.9)	122.51		
6	6.81	d (5.5)	117.6		
A	3.81, 4.02	m	72.5	1'	1'
B	2.76	t (6.8)	35.8		1'
Inner β -Glucose					
1'	4.46	d (8.2)	103.4	α	α
2'	3.36	m	76.9		1''
3'	3.77	m	82.0	1''	1''
4'	4.97	t (9.7)	70.1	C=O	
5'	3.75	m	74.4		1'''
6'	3.61, 3.95	m	69.1	1'''	1'''
Rhamnose					
1''	5.05	d (1.6)	102.9	3'	2', 3'
2''	3.59	m	71.2		
3''	3.96	m	71.6		
4''	3.25	m	73.7		
5''	3.42	m	70.9		
6''	0.98	d (6.2)	18.4		α' , β' , 2''''

Outern β -Glucose					
1'''	4.30	d (7.9)	103.9	6'	5', 6'
2'''	3.22	dd (9.1, 7.9)	74.2		
3'''	3.28	m	77.2		
4'''	3.25	m	70.8		
5'''	3.29	m	76.9		
6'''	3.60, 3.78	m	62.0		
Caffeoyl Ester					
1''''	\	\	128.0		
2''''	7.15	d (1.9)	116.6		6''
3''''	\	\	145.7		
4''''	\	\	149.0		
5''''	7.08	dd (8.4, 1.9)	124.3		
6''''	6.88	d (8.4)	117.6		
α'	6.33	d (15.8)	115.1		6''
β'	7.64	d (15.8)	149.0		6''
C=O	\	\	169.5	4'

Figure S1. TOCSY spectrum (110 ms mixing time) of Echinacoside (**2**)

Figure S2. HSQC spectrum of Echinacoside (**2**)

Figure S3. HMBC spectrum ($J=4\text{Hz}$) of Echinacoside (**2**)

Figure S4. NOESY spectrum (700 ms mixing time) of Echinacoside (**2**)