Kisiriko, M., Noleto-Dias, C., Bitchagno-Mbahbou, G., Naboulsi, I., Anastasiadi, M., Terry, L.A., Sobeh, M., Beale, M. H. and Ward, J. L. 2024. The First Comprehensive Chemical Profiling of Vachellia gummifera (Willd.) Kyal. & Boatwr., a Plant with Medicinal Value. *Chemistry & Biodiversity.* p. e202400396. https://doi.org/10.1002/cbdv.202400396

The First Comprehensive Chemical Profiling of *Vachellia gummifera* (Willd.) Kyal. & Boatwr., a Plant with Medicinal Value

Musa Kisiriko[a,b,c], Clarice Noleto-Dias[a], Gabin T.M. Bitchagno[a,c]‡, Imane Naboulsi[c], Maria Anastasiadi[b], Leon A. Terry[b], Mansour Sobeh[c], Michael H. Beale[a] and Jane L. Ward*[a].

[a] Plant Sciences for the Bioeconomy, Rothamsted Research, West Common, Harpenden AL5 2JQ, UK

[b] Plant Science Laboratory, Cranfield University, Cranfield, MK43 0AL, UK

[c] AgroBioSciences Research Division, Mohammed VI Polytechnic University, Lot 660, Moulay Rachid, Ben Guerir 43150, Morocco

‡ Current Address: Royal Botanical Gardens, Kew, Richmond, London, TW9 3AE

Corresponding author:

E-mail : jane.ward@rothamsted.ac.uk

Abstract

Vachellia gummifera (Willd.) Kyal. & Boatwr. is a medicinal plant endemic to Morocco that has no documented studies on its chemical composition. In this study, the chemical composition of the water/methanol (4:1) extracts of air-dried leaf and stem samples of Moroccan *V. gummifera* was determined using UHPLC-MS and NMR. In total, over 100 metabolites were identified in our study. Pinitol was the major compound in both the leaf and stem extracts, being significantly more abundant in the former. Asparagine and 3-hydroxyheteroendrin were the second most abundant compounds in the stem and leaf extracts, respectively, though both compounds were present in each tissue. The other compounds included flavonoids based on quercetin, and phenolic derivatives. Eucomic acid, only identified in the stems and was the major aromatic compound distinguishing the leaf and stem profiles. Quercetin 3-*O*-(6″-*O*-malonyl)-*β*-D-glucopyranoside was identified as the major flavonoid in the leaves but was also present in the stems. Other malonylated derivatives that were all flavonol glycosides based on myricetin, kaempferol, and isorhamnetin in addition to quercetin were also identified. This is the first report of eucomic acid and malonylated compounds in *Vachellia* species. This report provides valuable insights into the chemotaxonomic significance of the *Vachellia* genus.

Keywords

Vachellia gummifera, pinitol, eucomic acid, 3-hydroxyheteroendrin, flavonoids.

Introduction

Vachellia gummifera (Willd.) Kyal. & Boatwr. (Basionym: *Acacia gummifera* Willd.) is a thorny flowering plant belonging to the Fabaceae family.[1] It is endemic to Morocco where it is commonly referred to as *gommier du Maroc* and known by the vernacular name *Telh*.[2-3] In Morocco, the plant′s various parts are used in traditional medicine to treat different ailments. For instance, the aerial parts are used to treat bronchitis and cough,[4] decoctions from its roots

for type 2 diabetes,[5] poultice of the leaves for wounds and the powder of the bark for measles.[3] Its extracts also showed *in vitro* nematicidal activity against *Meloidogyne* ssp.[6] *V. gummifera* is an important forage plant for Moroccan Dorcas gazelles[7] and its powder biomass can be used as a biosorbent for removing lead and cadmium from polluted water.[8]

Like *Acacia sensu lato*, the genus from which the *Vachellia* sub genus was derived, the phytochemistry of many constituent species has not been studied yet.[9-10] To the best of our knowledge, there exists no reported studies on the chemical composition of *V. gummifera*. Mouhajir et al., (2001)[11] used electronic spin resonance spectroscopy, a method that detects phenolics with free ortho- or para-dihydroxy groups to study compounds in the plant. However, they did not detect any compounds in their study despite phenolics being some of the most abundant secondary metabolites in plants.[12]

Since this plant has reported apparent beneficial effects as elaborated above, it is important to assess its phytochemical composition which might provide valuable insights to understand the chemistry responsible for the effects. In line with the above, the objective of this study was to characterise the phytochemical profile of the polar extracts of *V. gummifera*. The leaves and stems were investigated in this study as they make up the major biomass on the plant and are prominently used in traditional medicine and other applications. This is the first report on the characterisation of this plant′s metabolome.

Results and Discussion

The UHPLC-MS total ion chromatograms (Figure 1) and the ^{1}H NMR spectra (Figure 2) of the leaf and stem polar extracts showed that many of the compound peaks were common to both extracts. However, certain peaks were also observed to be unique to each extract. To identify the compounds responsible for these peaks, each extract was fractionated by HPLC, and the resultant fractions were analysed separately by UHPLC-MS/MS and NMR.

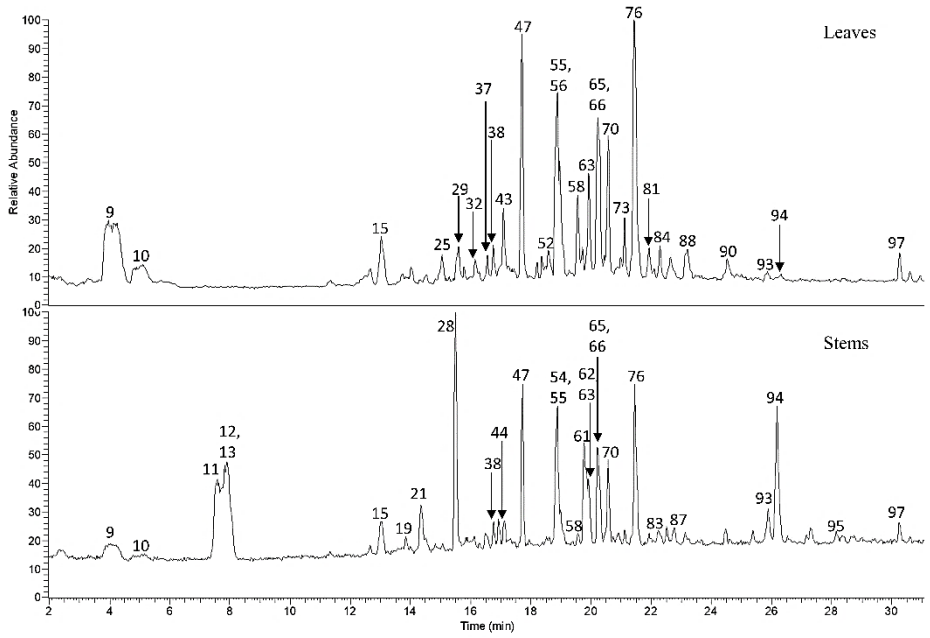


Fig. 1. Total ion chromatograms (negative ion mode) of the leaves and stems of *V. gummifera* extracted with H_2O:CH_3OH (4:1 v/v)

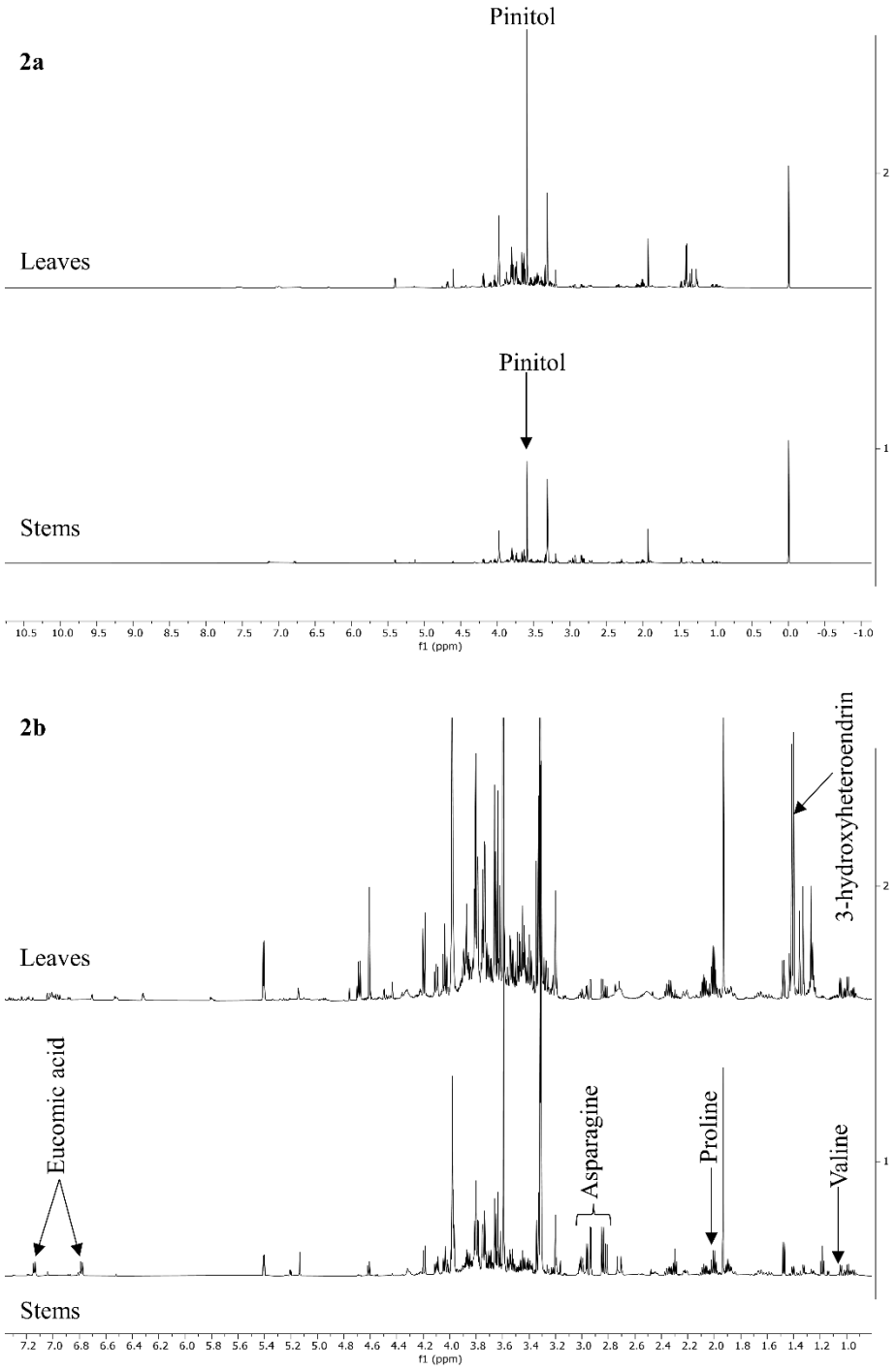


Fig. 2. NMR traces of *V. gummifera* leaf and stem crude extracts recorded at 600MHz in D_2O:CD_3OD (4:1 v/v) referenced to TSP-d_4, 0.01% w/v at δ 0.00

Known compounds were identified by comparing their data to our in-house database or to the literature. Putative identification of unknown compounds was carried out using UHPLC-MS accurate masses and MS/MS fragmentation patterns. Identities were confirmed via isolation and structural elucidation by 2D-NMR wherever possible. Table 1 summarises the UHPLC-MS data and the level of identification of the isolated compounds in both the leaf and stem extracts. The chemical structures of selected compounds from both extracts that were confirmed by NMR or using a standard are shown in Figure 3.

Table 1: Compound profile of the leaf and stem extracts of *V. gummifera* in water:methanol (4:1)

Compound No.	**R_t** (min)	**[*M*-H]$^-$** (*m/z*)	**Molecular Formula**	**Δ** (ppm)	**MS/MS product ions** (*m/z*)	**Identity**	**Extract**
1	1.01	207.0148	$C_6H_8O_8$	3.59	127.0042, 133.0148, **189.0045**	Hydroxycitric acid [c]	Leaves
2	1.03	189.0043	$C_6H_6O_7$	4.03	**127.0042**, 170.9944	Hibiscus acid [c]	Leaves, stems
3	1.13	133.0152	$C_4H_6O_5$	7.44	71.0144, 89.0251, **115.0008**	Malic acid [a]	Leaves
4	1.50	191.0199	$C_6H_8O_7$	3.96	**111.0093**, 129.0200, 173.0098	Citric acid [a]	Leaves, stems
5	1.75	117.0201	$C_4H_6O_4$	6.64	**73.0302**, 99.0095	Succinic acid [a]	Stems
6	2.18	331.067	$C_{13}H_{16}O_{10}$	1.51	151.0043, **169.0148**, 211.0253, 241.0359, 271.0464	*β*-Glucogallin (1-*O*-galloyl-*β*-D-glucopyranoside) [b]	Leaves, stems
7	2.42	169.0146	$C_7H_6O_5$	2.11	**125.0249**	Gallic acid [a]	Leaves, stems
8	3.22	297.1191*	$C_{10}H_{20}O_7$	1.85	101.0249, 159.0305, 161.0458, **251.1137**	Alkyl alcohol glucoside [c]	Leaves
9	3.96	322.1143*	$C_{11}H_{19}NO_7$	1.41	159.0667, **161.0460**, 188.0570, 218.0677, 249.0982	3-hydroxyheteroendrin [b]	Leaves, stems
10	5.12	295.1034*	$C_{10}H_{18}O_7$	3.56	101.0249, 159.0670, **161.0463**, 173.9727, 249.0984	Alkyl glucoside [c]	Leaves, stems
11	7.64	255.0507	$C_{11}H_{12}O_7$	-1.20	**165.0562**, 179.0354, 193.0510	Piscidic acid isomer 1 [b]	Stems
12	7.72	255.0504	$C_{11}H_{12}O_7$	-2.29	**165.0562**, 179.0354, 193.0510	Piscidic acid isomer 2 [b]	Stems
13	7.77	255.0517	$C_{11}H_{12}O_7$	-2.69	**165.0565**, 179.0358, 193.0514	Piscidic acid isomer 3 [b]	Stems
14	11.39	203.0828	$C_{11}H_{12}N_2O_2$	6.24	74.0252, **116.0510**, 142.0667, 159.0932, 186.0564	Tryptophan [a]	Leaves, stems
15	13.09	183.0300	$C_8H_8O_5$	0.68	124.0170, **168.0073**	Methyl gallate [b]	Leaves, stems
16	13.14	371.0984*	$C_{15}H_{18}O_8$	-0.01	119.0508, **163.0404**, 325.0930	(*E*)-*p*-Coumaric acid 4-*O*-*β*-D-glucopyranoside [b]	Leaves
17	13.2	451.2183*	$C_{19}H_{34}O_9$	0.77	101.0248, 161.0459, **243.1602**, 405.2130	Unknown	Leaves
18	13.68	577.1343	$C_{30}H_{26}O_{12}$	-1.39	125.0248, **289.0716**, 407.0767	Procyanidin B3 [b]	Stems
19	13.87	305.0665	$C_{15}H_{14}O_7$	1.32	**125.0251**, 165.0199, 179.0355, 219.0668, 261.0773	(Epi)gallocatechin [c]	Leaves, stems
20	14.04	359.0983	$C_{15}H_{20}O_{10}$	-0.10	**197.0458**, 211.0616, 239.0563, 299.0775	Syringic acid *O*-*β*-D-glucopyranosyl ester (Erigeside C) [b]	Leaves, stems
21	14.37	289.0719	$C_{15}H_{14}O_6$	-1.11	125.0251, 179.0356, 205.0512, **245.0824**	Catechin [a]	Stems

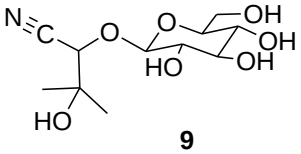
22	14.53	761.1347	$C_{37}H_{30}O_{18}$	-1.64	125.0247, 169.0149, 177.0194, 305.0662, **423.0712**	(Epi)gallocatechin-(epi)gallocatechin gallate [c]	Leaves, stems
23	14.84	465.1035	$C_{21}H_{22}O_{12}$	-0.72	275.0567, 285.0410, **303.0516**, 343.0678	Taxifolin 7-*O*-*β*-D-glucopyranoside [b]	Stems
24	14.88	457.1349	$C_{20}H_{26}O_{12}$	-0.64	**163.0405**, 205.0511, 325.0922	*p*-Coumaric acid pentosyl hexoside [c]	Stems
25	15.10	257.1396	$C_{13}H_{22}O_5$	0.56	101.0248, 139.0769, 155.1080, 195.1392, **213.1497**	Unknown	Leaves, stems
26	15.33	449.2031*	$C_{19}H_{32}O_9$	0.56	161.0463, 223.1340, **241.1452**, 403.1995	Megastigman-7-ene-6,9,10-triol-3-one 9-*O*-*β*-D-glucopyranoside [b]	Leaves
27	15.35	771.1978	$C_{33}H_{40}O_{21}$	-1.53	299.0191, 300.0263, 301.0345, **462.0785**, 609.1440	Quercetin *O*-rutinoside *O*-hexoside [c]	Leaves
28	15.50	239.0557	$C_{11}H_{12}O_6$	-1.70	149.0614, 177.0563, **179.0354**, 195.0667, 221.0461	Eucomic acid [b]	Stems
29	15.56	325.0931	$C_{15}H_{18}O_8$	4.14	**163.0404**	(*Z*)-*p*-Coumaric acid 4-*O*-*β*-D-glucopyranoside [b]	Leaves, stems
30	15.82	417.1401*	$C_{17}H_{24}O_9$	-0.39	194.0585, **209.0822**	Syringin [b]	Stems
31	16.10	327.1085	$C_{15}H_{20}O_8$	-0.05	123.0456, **165.0560,** 267.0887	Phenyl hexoside derivative [c]	Stems
32	16.18	447.1870*	$C_{19}H_{30}O_9$	-0.50	161.0458, **221.1183**, 401.1810	Unknown	Leaves
33	16.34	771.1977	$C_{33}H_{40}O_{21}$	-0.86	301.0352, **462.0798**, 609.1451	Quercetin 3-*O*-rutinoside-7-*O*-*β*-D-glucopyranoside [b]	Leaves
34	16.49	477.1612	$C_{20}H_{30}O_{13}$	-0.49	89.0248, 125.0249, 183.0665, **233.0668**, 293.0879	3,4,5-trimethoxyphenol 1-*O*-*β*-D-apiofuranosyl-(1→6)-*β*-D-glucopyranoside (Kelampayoside A) [b]	Stems
35	16.52	289.0718	$C_{15}H_{14}O_6$	1.96	179.0357, 205.0512, **245.0824**	Epicatechin [a]	Stems
36	16.55	385.1140	$C_{17}H_{22}O_{10}$	-0.09	206.0586, **223.0609**	(*Z*)-Sinapic acid *O*-*β*-D-glucopyranoside [b]	Leaves
37	16.58	447.1143*	$C_{17}H_{23}O_{11}$	-0.29	113.0251, **197.0462**, 267.0730, 271.0833, 429.1046	Syringic acid derivative [c]	Leaves
38	16.76	457.1349	$C_{20}H_{26}O_{12}$	-0.42	119.0506, **163.0402**, 205.0507, 325.0927	(*Z*)-*p*-Coumaric acid 4-*O*-(2′-*O*-*β*-D-apiofuranosyl)-*β*-D-glucopyranoside [b]	Leaves, stems
39	16.78	595.1661	$C_{27}H_{32}O_{15}$	-1.29	355.0821, **385.0926**, 415.1031, 475.1241	Unknown	Leaves
40	16.93	589.1767	$C_{25}H_{34}O_{16}$	-1.17	163.0403, 325.0929, **457.1341**	*p*-Coumaric acid dipentosyl hexoside [c]	Leaves
41	17.03	711.1409	$C_{30}H_{32}O_{20}$	-0.48	299.0205, 301.0358, **462.1045**, 463.1172, 505.1366, 625.1419, 667.1522	Quercetin 3-*O*-(6″-*O*-malonyl)-*β*-D-glucopyranosyl-7-*O*-*β*-D-glucopyranoside [b]	Leaves
42	17.10	428.1193	$C_{18}H_{23}NO_{11}$	0.1	151.0044, 168.0071, 178.0152, 248.0570, 313.0571, **401.1093**	Cyanogenic derivative [c]	Leaves

43	17.11	431.1921	$C_{20}H_{32}O_{10}$	0.80	101.0248, 113.0248, **119.0354**, 161.0459, 179.0563	Unknown	Leaves
44	17.12	457.0774	$C_{22}H_{18}O_{11}$	-0.42	125.0247, 161.0245, **169.0141**, 305.0662, 331.0452	Epigallocatechin gallate [b]	Leaves, stems
45	17.22	329.0878	$C_{14}H_{18}O_{9}$	-0.02	**167.0353**, 191.0353, 209.0459	Vanillic acid hexoside [c]	Leaves, stems
46	17.38	437.2391*	$C_{19}H_{36}O_{8}$	0.92	161.0460, 229.1816, **391.2338**	Unknown	Leaves
47	17.74	593.1499	$C_{27}H_{30}O_{15}$	-2.20	353.0672, 383.0779, **473.1093**, 503.1203	Apigenin 6,8-di-*C*-*β*-D-glucopyranoside [b]	Leaves, stems
48	17.91	561.1397	$C_{30}H_{26}O_{11}$	-0.93	161.0244, 245.0813, 271.0605, **289.0711**, 391.0821	B-type proanthocyanidin [c]	Stems
49	18.01	563.1402	$C_{26}H_{28}O_{14}$	-0.77	353.0662, 383.0768, **443.0976**, 473.1082	Apigenin 6,8-di-*C*-pentosyl hexoside [c]	Stems
50	18.12	465.1036	$C_{21}H_{22}O_{12}$	-0.52	125.0249, 177.0196, 259.0613, **285.0404**, 303.0509	Taxifolin hexoside [c]	Stems
51	18.39	565.0831	$C_{24}H_{22}O_{16}$	-0.72	**316.0222**, 317.0300, 271.0249, 479.0826, 521.0935	Myricetin 3-*O*-malonyl hexoside [c]	Leaves
52	18.61	479.0829	$C_{21}H_{20}O_{13}$	-0.32	**316.0220**, 317.0290	Myricetin 3-*O*-*β*-D-glucopyranoside [a]	Leaves, stems
53	18.71	625.1402	$C_{27}H_{30}O_{17}$	-1.32	178.9989, **316.0223**, 317.0297	Myricetin 3-*O*-rutinoside [a]	Stems
54	18.84	479.0828	$C_{21}H_{20}O_{13}$	-0.67	**316.0219**, 317.0296	Myricetin 3-*O*-hexoside [c]	Stems
55	18.87	563.139	$C_{26}H_{28}O_{14}$	-2.83	353.0679, 383.0786, **443.0996**, 473.1104, 545.1321	Apigenin (6-*C*-*α*-L-arabinopyranosyl)-8-*C*-*β*-D-glucopyranoside [b]	Leaves, stems
56	18.99	755.2028	$C_{33}H_{40}O_{20}$	1.46	271.0252, **300.0277**, 301.0327, 489.1049, 591.1353	Quercetin 3-*O*-di-*α*-L-rhamnopyranosyl-*β*-D-glucopyranoside [b]	Leaves, stems
57	19.2	433.114	$C_{21}H_{22}O_{10}$	-0.09	271.0614, **313.0718**, 343.0824	Naringenin *C*-hexoside [c]	Stems
58	19.55	609.1453	$C_{27}H_{30}O_{16}$	-1.39	**300.0281**, 301.0332, 445.0770	Quercetin 3-*O*-*α*-L-rhamnopyranosyl-(1→2)-*β*-D-galactopyranoside [b]	Leaves, stems
59	19.60	565.083	$C_{24}H_{22}O_{16}$	-0.95	178.9989, **316.0220,** 317.0299, 339.0120, 521.0933	Myricetin 3-*O*-(6″-*O*-malonyl)-*β*-D-glucopyranoside [b]	Leaves, stems
60	19.72	609.1453	$C_{27}H_{30}O_{16}$	-1.29	**300.0281**, 301.0332, 445.0770	Quercetin 3-*O*-*α*-L-rhamnopyranosyl-(1→6)-*β*-D-galactopyranoside [b]	Leaves
61	19.75	493.1345	$C_{23}H_{26}O_{12}$	-1.27	303.0876, **331.0822**	Di-*O*-methyltaxifolin-*β*-D-glucopyranoside [b]	Stems
62	19.81	303.0506	$C_{15}H_{12}O_{7}$	-1.51	125.0250, 177.0199, **285.0408**	Taxifolin [a]	Stems
63	19.96	609.145	$C_{27}H_{30}O_{16}$	-1.74	**300.0271**, 301.0332, 457.0794, 591.1362	Quercetin 3-*O*-neohesperidoside [b]	Leaves, stems
64	20.13	431.0983	$C_{21}H_{20}O_{10}$	0.95	283.0613, **311.0561**, 341.0669	Apigenin *C*-hexoside [c]	Stems

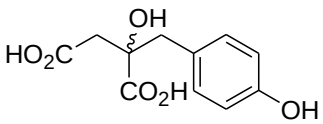
65	20.23	609.1453	$C_{27}H_{30}O_{16}$	-1.39	**300.0277**, 301.0346	Rutin [a]	Leaves, stems
66	20.30	463.0875	$C_{21}H_{20}O_{12}$	-1.63	**300.0279**, 301.0356	Quercetin 3-*O*-*β*-D-galactopyranoside [b]	Leaves, stems
67	20.44	769.2185	$C_{34}H_{42}O_{20}$	-0.77	178.9993, 271.0258, 299.0204, **314.0436**, 315.0509, 339.0529, 605.1527	Isorhamnetin 3-*O*-dirhamnosylhexoside [c]	Stems
68	20.44	303.0509	$C_{15}H_{12}O_{7}$	-0.57	125.0245, 177.0191, 217.0503, 275.0554, **285.0397**	Epitaxifolin [b]	Stems
69	20.45	595.1662	$C_{27}H_{32}O_{15}$	-1.13	135.0459, 151.0044, 175.0044, **287.0567**	Eriocitrin [a]	Leaves
70	20.57	463.0875	$C_{21}H_{20}O_{12}$	-1.56	**300.0281**, 301.0359, 445.0733	Quercetin 3-*O*-*β*-D-glucopyranoside [b]	Leaves, stems
71	20.65	137.0248	$C_{7}H_{6}O_{3}$	6.92	**93.0351**	Salicylic acid [a]	Stems
72	20.9	447.0928	$C_{21}H_{20}O_{11}$	0.11	284.0332, **285.0407**	Luteolin 7-*O*-*β*-D-glucopyranoside [a]	Stems
73	21.11	549.0882	$C_{24}H_{22}O_{15}$	-0.66	**300.0269**, 301.0348, 323.0172, 463.0871, 505.0987	Quercetin 3-*O*-(6″-*O*-malonyl)-*β*-D-galactopyranoside [b]	Leaves, stems
74	21.12	597.1819	$C_{27}H_{34}O_{15}$	-0.01	315.0874, **357.0977**, 387.1082, 417.1187, 477.1400	Phloretin 3′,5′-di-*C*-*β*-D-glucopyranoside [b]	Stems
75	21.34	535.1088	$C_{24}H_{24}O_{14}$	-1.05	287.0204, 315.0152, **330.0385**, 331.0463, 475.0883, 493.1001	Hydroxy-methoxyquercetin acetyl hexoside [c]	Stems
76	21.46	549.0882	$C_{24}H_{22}O_{15}$	-0.70	**300.0280**, 301.0358, 323.0168, 463.0888, 505.0981	Quercetin 3-*O*-(6″-*O*-malonyl)-*β*-D-glucopyranoside [b]	Leaves, stems
77	21.49	493.1348	$C_{23}H_{26}O_{12}$	-0.72	165.0193, 303.0869, **331.0814**, 373.0919	Hydroxy-methoxyquercetin hexoside [c]	Stems
78	21.51	549.0882	$C_{24}H_{22}O_{15}$	-0.53	**300.0266**, 301.0343, 323.0169, 463.0868, 505.0987	Quercetin 3-*O*-malonyl-hexoside [c]	Leaves
79	21.67	623.1608	$C_{28}H_{32}O_{16}$	-0.69	299.0203, **314.0433**, 315.0511	Isorhamnetin 3-*O*-rhamnosyl hexoside [c]	Stems
80	21.93	623.1608	$C_{28}H_{32}O_{16}$	-1.49	151.0041, 271.0245, 300.0272, 313.0361, **315.0508**	Isorhamnetin rutinoside [c]	Stems
81	22.07	447.0925	$C_{21}H_{20}O_{11}$	0.66	**284.0324**, 285.0402	Kaempferol 3-*O*-*β*-D-glucopyranoside [b]	Leaves
82	22.08	519.1869	$C_{26}H_{32}O_{11}$	0.58	151.0405, 311.1285, 342.1105, **357.1344**	Pinoresinol hexoside [c]	Stems
83	22.19	287.056	$C_{15}H_{12}O_{6}$	-0.28	125.0250, 180.0070, 201.0562, 243.0667, **259.0616**	Aromadendrin [a]	Stems
84	22.31	549.0882	$C_{24}H_{22}O_{15}$	-0.65	**300.0270**, 301.0342, 463.0875, 505.0985	Quercetin 3-*O*-malonyl-hexoside [c]	Leaves
85	22.36	477.1035	$C_{22}H_{22}O_{12}$	0.38	**314.0437**, 315.0517, 357.0625	Isorhamnetin 3-*O*-*β*-D-glucopyranoside [a]	Stems
86	22.58	431.0982	$C_{21}H_{20}O_{10}$	0.81	**268.0373**, 269.0451	Apigenin 7-*O*-*β*-D-glucopyranoside [a]	Stems

87	23.08	317.0667	$C_{16}H_{14}O_7$	1.95	125.0251, 152.0121, 165.0572, 180.0071, 231.0669, 273.0776, **289.0724**	Methyltaxifolin [c]	Stems
88	23.13	533.0934	$C_{24}H_{22}O_{14}$	-0.49	**284.0331**, 285.0406	Kaempferol 3-*O*-(6″-*O*-malonyl)-*β*-D-glucopyranoside [b]	Leaves
89	23.58	563.1038	$C_{25}H_{24}O_{15}$	-0.85	299.0200, **314.0432**, 315.0510	Isorhamnetin 3-*O*-malonyl hexoside [c]	Leaves, stems
90	24.58	553.1898	$C_{26}H_{34}O_{13}$	-5.19	271.0950, **289.1054**, 421.1471	Unknown	Leaves
91	25.70	491.1193	$C_{23}H_{24}O_{12}$	1.80	314.0433, **329.0665**, 476.0953	Tricin hexoside [c]	Stems
92	25.88	301.0354	$C_{15}H_{10}O_7$	-0.22	178.9989, 151.0041	Quercetin [a]	Leaves
93	25.89	285.0406	$C_{15}H_{10}O_6$	0.38	241.0517	Luteolin [a]	Leaves, stems
94	26.15	331.082	$C_{17}H_{16}O_7$	-0.98	125.0251, **152.0122**, 180.0072, 316.0595	Di-*O*-methyltaxifolin [b]	Leaves, stems
95	28.09	271.061	$C_{15}H_{12}O_5$	-0.71	93.0351, 119.0508, **151.0042**, 169.0148, 177.0198	Naringenin [a]	Stems
96	28.37	269.0455	$C_{15}H_{10}O_5$	-0.01	**151.004**, 201.0565, 225.0572, 252.0449	Apigenin [a]	Stems
97	30.24	327.2177	$C_{18}H_{32}O_5$	-0.05	**171.1034**, 211.1347, 229.1453, 291.1974, 309.2082	Trihydroxy−octadecadienoic acid [c]	Leaves, stems
98	33.00	285.2073	$C_{16}H_{30}O_4$	0.66	**267.1962**	Hexadecanedioic acid [c]	Leaves
99	33.33	309.2072	$C_{18}H_{30}O_4$	0.21	171.1033, 185.1190, 251.1658, **291.1971**	Hydroperoxy-octadecatrienoic acid [c]	Leaves
100	36.76	311.2228	$C_{18}H_{32}O_4$	0.11	87.0457, 201.1138, **223.1710**, 235.1709, 275.2023, 293.2129	Hydroperoxy-octadecadienoic acid [c]	Leaves
101	39.67	293.2123	$C_{18}H_{30}O_3$	0.45	183.1400, 235.1712, **275.2027**	Oxo-octadecadienoic acid (isomer 1) [c]	Leaves
102	39.84	293.2125	$C_{18}H_{30}O_3$	0.88	171.1034, 183.1398, 195.1400, 211.1346, 235.1711, **275.2024**	Oxo-octadecadienoic acid (isomer 2) [c]	Leaves

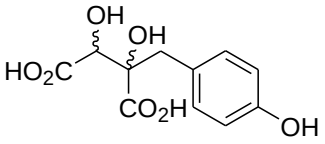
* Corresponds to the formate adduct. [a] Identification based on comparison with standards or match in our database. [b] Identification based on isolation of the compounds and NMR. [c] Putative identification based on UHPLC-MS data. Product ions in bold are the base peaks of the MS/MS spectra. R_t is the retention time.



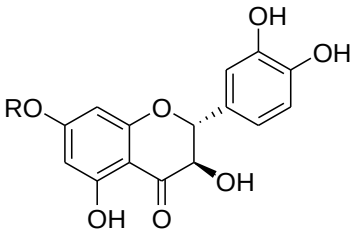
9



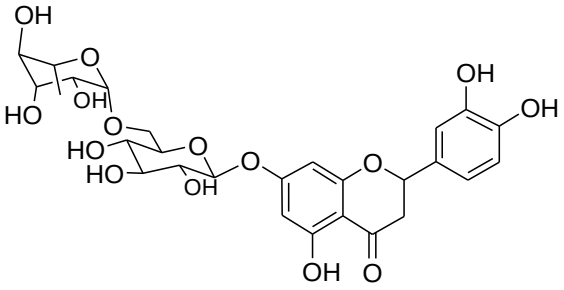
28



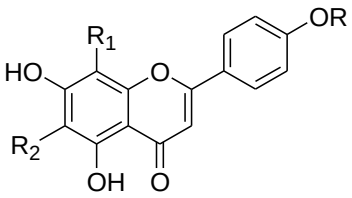
11-**13**



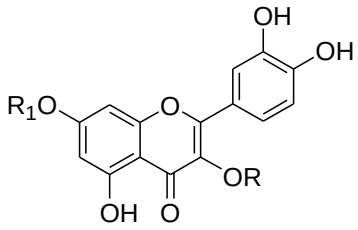
62 R = H
68 R = H
23 R = glc



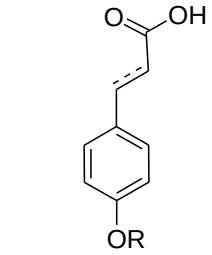
69



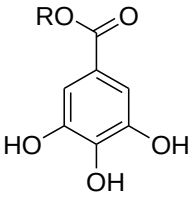
	R	R_1	R_2
47	H	glc	glc
55	H	glc	ara
86	glc	H	H
96	H	H	H



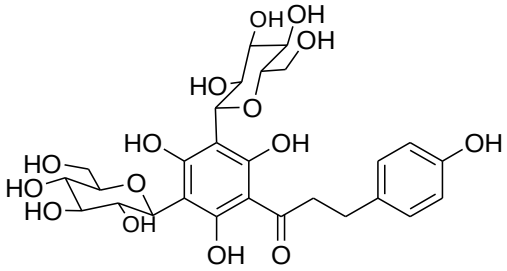
	R	R_1
33	rut	glc
41	A	glc
56	B	H
58	C	H
63	neo	H
65	rut	H
66	gal	H
70	glc	H
73	D	H
76	A	H



16 R = glc
29 R = glc
38 R = E



6 R= glc
15 R= CH_3



74

Fig. 3. Chemical Structures of selected compounds in the leaf and stem extracts of *V. gummifera*. Abbreviations – glc = *β*-D-glucopyranose; gal = *β*-D-galactopyranose; ara = *α*-L-arabinopyranose; rut = rutinose; neo = neohesperidose; A = (6″-*O*-malonyl)-*β*-D-glucopyranose; B = di-*α*-L-rhamnopyranosyl-*β*-D-glucopyranose; C = *α*-L-rhamnopyranosyl-(1→2)-*β*-D-galactopyranose; D = (6″-*O*-malonyl)-*β*-D-galactopyranose; E = (2′-*O*-*β*-D-apiofuranosyl)-*β*-D-glucopyranose.

Major compounds and differences between the leaf and stem profiles.

Inspection of the ^{1}H NMR spectra of both the leaf and stem crude extracts revealed several differences in the distribution of compounds found in both plant parts. The major peak in both spectra is a singlet at δ_H 3.59 (s, 3H) (Figure 2a). Additional signals at δ_H 3.64 (t, J = 9.8 Hz, 1H), 3.74 (dd, J = 9.9, 2.8 Hz, 1H), 3.80 (dd, J = 9.8, 2.8 Hz, 1H), 3.98 (m), and another overlapping with the residual methanol signal corresponded to pinitol when compared with the ^{1}H NMR data of an authentic standard of the compound. Quantification of the compound in both crude extracts by qNMR revealed that it was significantly more abundant in the leaves with 64.83 ± 0.88 mg g^{-1} DW compared to the stems where 24.89 ± 0.06 mg g^{-1} DW was observed (Figure 4). Pinitol has previously been identified in *Vachellia* species including *V. nilotica*[13], *V. farnesiana*[14], and *V. etbaica* Schweinf.[15] as well as several *Acacia* species.[9-10] In addition to various pharmacological effects including hepatoprotective, anticancer, cardioprotective and anti-inflammatory effects, several studies have reported the anti-diabetic properties of the compound.[16-17] This probably explains the reported use of *V. gummifera* in traditional medicine for treating type 2 diabetes. Pinitol is also a known osmoprotective compound that accumulates in plants in response to water and salinity stress.[18-19] Therefore, the presence of the compound as the major metabolite in the extracts could be a consequence of the plant material having been sampled from plants growing in Morocco which is an arid area implying that they experienced significant drought and water stress during their growth.

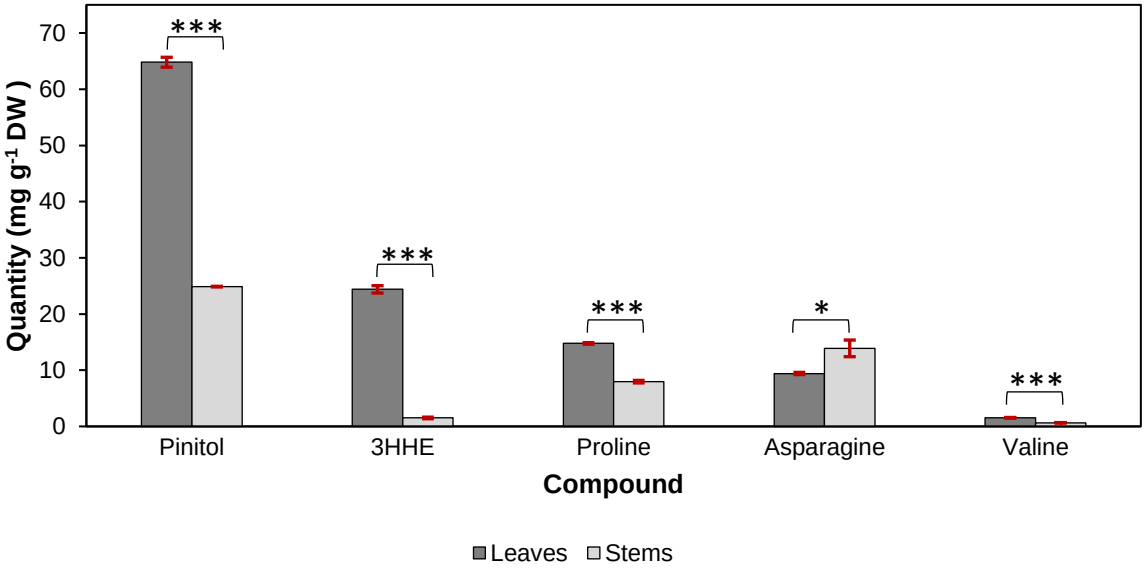


Fig. 4. Comparison of the quantities of selected compounds from leaves and stems of *V. gummifera* using the *t*-test. Bars represent the mean values, and the error bars represent the standard deviation. 3HHE – 3-Hydroxyheteroendrin. *p < 0.05 and ***p < 0.001.

The upfield region of the ^{1}H NMR spectra of the crude extracts also showed the presence of various clear signals arising from amino acids (Figure 2b). The identity of these amino acids was confirmed by comparing their multiplet peaks in the crude spectra with authentic standards. The presence of asparagine was confirmed from its characteristic signals at δ_H 2.83 (dd, J = 16.9, 8.0 Hz, 1H) and 2.95 (dd, J = 16.9, 4.1 Hz, 1H). Asparagine was significantly more abundant in the stems (13.90 ± 1.50 mg g^{-1} DW) compared to the leaves (9.41 ± 0.20 mg

g^{-1} DW), and in fact a careful examination of the whole spectrum reveals that it is the second major compound after pinitol in the stem extract. Proline, which was significantly more abundant in the leaves (14.79 ± 0.13 mg g^{-1} DW) compared to the stems (7.95 ± 0.23 mg g^{-1} DW) was also confirmed from some of its signals at δ_H 2.00 (m, 2H), 2.07 (m, 1H), 2.34 (m, 1H) and 4.10 (dd, *J* = 8.8, 6.4 Hz, 1H). Just like pinitol, proline is also known to accumulate in plants that experience water and salt stress.[18-19] Additionally, valine with 1.54 ± 0.03 mg g^{-1} DW in the leaves and 0.65 ± 0.01 mg g^{-1} DW in the stems was also confirmed from its characteristic signals at δ_H 0.99 (d, *J* = 7.0 Hz, 3H) and 1.04 (d, *J* = 7.0 Hz, 3H).

However, the main difference in the upfield region of the spectra of both extracts was the levels of two aliphatic singlets at δ_H 1.40 and 1.39 ppm. In fact, these signals represented the second most abundant compound in the leaves, after pinitol. Upon HPLC fractionation of both extracts, those singlets together with an additional singlet at *ca.* δ_H 4.60 ppm corresponded to compound **9**. In UHPLC-MS (negative ionisation mode), this compound showed a molecular ion at *m/z* 322.1143 corresponding to the formate adduct, $[M+HCOOH-H]^-$, of a compound with formula $C_{11}H_{19}NO_7$. The MS/MS of the parent ion showed a fragment at *m/z* 249.0982 ($C_{10}H_{17}O_7^-$) resulting from the loss of HCN indicating the presence of a nitrile group and another at *m/z* 161.0460 ($C_6H_9O_5^-$) corresponding to a hexose fragment. Comparison of its ^{1}H NMR data to that of a compound previously isolated from *V. sieberiana* var. woodii confirmed the identity of **9** to be 3-hydroxyheteroendrin.[20] This compound was significantly more abundant in the leaves with 24.40 ± 0.65 mg g^{-1} DW compared to the stems with just 1.52 ± 0.14 mg g^{-1} DW.

On the other hand, inspection of the aromatic region of the ^{1}H NMR spectra of both the leaf and stem extracts revealed two major signals at δ_H 7.14 (d, *J* = 8.4 Hz) and 6.78 (d, *J* = 8.4 Hz) that were present in the stems but absent in the leaves (Figure 2b). These signals represented the major difference in the aromatic profile of the two extracts. From the isolated fractions, both doublets corresponded to compound **28**. This compound eluted at rt 15.50 min, and its MS showed a deprotonated molecular ion ($[M-H]^-$), at *m/z* 239.0557 corresponding to a compound with molecular formula $C_{11}H_{12}O_6$. The MS/MS of the parent ion produced fragment ions at *m/z* 149.0614 ($C_9H_9O_2^-$), 177.0563 ($C_{10}H_9O_3^-$), 179.0354 ($C_9H_7O_4^-$), 195.0667 ($C_{10}H_{11}O_4^-$) and 221.0461 ($C_{11}H_9O_5^-$). The ^{1}H NMR spectrum of the isolated fraction showed two aromatic signals similar to those in the crude extract spectrum at δ_H 7.12 (d, *J* = 8.5 Hz, 2H) and 6.80 (d, *J* = 8.5 Hz, 2H) indicating a para-substitution of a phenyloxy, as well as two pairs of aliphatic methylene signals at δ_H 2.97 (d, *J* = 13.8 Hz, 1H)/ 2.86 (d, *J* = 13.8 Hz, 1H) and 2.93 (d, *J* = 16.1 Hz, 1H)/ 2.66 (d, *J* = 16.1 Hz, 1H). This data was consistent with that of eucomic acid when compared to reported LC-MS and NMR data in the literature.[21-22] To the best of our knowledge, this is the first time eucomic acid has been reported in any *Vachellia* or *Acacia* species. Unlike the spectra of the stem extracts, the aromatic region in the ^{1}H NMR spectra of the leaf extracts was dominated by flavonoid related signals, many of which were also present in the spectra from stem extracts, albeit in differing abundances.

Interestingly, inspection of the total ion chromatograms of the leaves and stems showed that the $[M-H]^-$ peak of eucomic acid (**28**) at *m/z* 239.0557 was also the most abundant peak in the stem extract and absent in the leaf extract (Figure 1). Therefore, this compound can be considered as the main aromatic compound differentiating the UHPLC-MS profiles of the leaves and stems of *V. gummifera*. In addition to **28**, the compounds **11**, **12**, and **13** appeared only in the chromatogram of the stem extract and were not visible in extracts from the leaves (Figure 1). Compound **11** eluted at rt 7.64 min and its MS showed an $[M-H]^-$ ion at *m/z*

255.0507 corresponding to a compound with the molecular formula $C_{11}H_{12}O_7$. The MS/MS of the parent ion afforded product ions at *m/z* 165.0562 ($C_9H_9O_3^-$), 179.0354 ($C_9H_7O_4^-$) and 193.0510 ($C_{10}H_9O_4^-$). The ^{1}H NMR spectrum of the compound showed a similar aromatic pattern as **28** with signals at δ_H 7.13 (d, *J* = 8.5 Hz, 2H) and 6.78 (d, *J* = 8.5 Hz, 2H). However, its spectrum showed two aliphatic signals at δ_H 4.45 (s, 1H) and 3.02 (m, 2H). The data was consistent with that of piscidic acid when compared to reported LC-MS and NMR data in the literature.[23-24] Interestingly, the compounds that eluted at rt 7.72 (**12**) and 7.77 (**13**) min exhibited similar [M-H]$^-$ ions, MS/MS fragmentation patterns and ^{1}H NMR chemical shifts as compound **11**. They might represent two of the three other stereoisomers as expected for a two-stereocenter molecule like piscidic acid. This finding is even more important as stereochemical occurrences of secondary metabolites and their origin in the natural world remain a much discussed topic in natural product chemistry.[25] Recently, piscidic acid was identified as the major compound in the branch extracts of *V. nilotica* yet a significantly smaller amount was identified in its gum.[26]

Compound **94,** one of the major abundant peaks in the chromatogram of the stem extract, was found only at trace level in the leaf extract chromatogram. It showed the [M-H]$^-$ ion at *m/z* 331.082 corresponding to a compound with the molecular formula $C_{17}H_{16}O_7$. Analysis of its UHPLC-MS and ^{1}H NMR data (Table 2) showed that it agreed with the identity of di-*O*-methyltaxifolin.[27] However, the data was not sufficient for confirmation of the exact methylation positions.

Table 2: ^{1}H NMR data of selected compounds from the leaves and stems of *V. gummifera*.

Compound Number	Identity	Extract	^{1}H NMR data
6	*β*-Glucogallin	L, S	δ_H 7.23 (s, 2H), 5.70 (d, *J* = 7.9 Hz, 1H), 3.90 (dd, *J* = 12.3, 2.1 Hz, 1H), 3.72 (dd, *J* = 12.3, 5.3 Hz, 1H), 3.48 – 3.26 (overlapping)
7	Gallic acid	L, S	δ_H 7.10 (s, 2H)
9	3-Hydroxyheteroendrin	L, S	δ_H 4.67 (d, *J* = 7.9 Hz, 1H), 4.59 (s, 1H), 3.87 (dd, *J* = 12.4, 2.2 Hz, 1H), 3.69 (dd, *J* = 12.4, 5.8 Hz, 1H), 3.47 (t, *J* = 9.4 Hz, 1H), 3.45 – 3.41 (m, 1H), 3.38 (d, *J* = 9.4 Hz, 1H), 3.26 (dd, *J* = 9.4, 7.9 Hz, 1H), 1.40 (s, 3H), 1.39 (s, 3H).
11	Piscidic acid (isomer 1)	S	δ_H 7.13 (d, *J* = 8.5 Hz, 2H), 6.78 (d, *J* = 8.5 Hz, 2H), 4.45 (s, 1H), 3.02 (m, 2H).
12	Piscidic acid (isomer 2)	S	δ_H 7.13 (d, *J* = 8.6 Hz, 2H), 6.79 (d, *J* = 8.6 Hz, 2H), 4.52 (s, 1H), 3.06 (d, *J* = 13.9 Hz, 1H), 3.03 (d, *J* = 13.9 Hz, 1H).
13	Piscidic acid (isomer 3)	S	δ_H 7.13 (d, *J* = 8.5 Hz, 2H), 6.78 (d, *J* = 8.5 Hz, 2H), 4.45 (s, 1H), 3.02 (m, 2H).
14	Tryptophan	L, S	δ_H 7.72 (d, *J* = 8.1 Hz, 1H), 7.51 (d, *J* = 8.1 Hz, 1H), 7.29 (s, 1H), 7.25 (ddd, *J* = 8.1, 7.1, 1.1 Hz, 1H), 7.17 (ddd, *J* = 8.1, 7.1, 1.0 Hz, 1H), 4.02 (dd, *J* = 8.3, 4.7 Hz, 1H), 3.48 (dd, *J* = 15.4, 4.7 Hz, 1H), 3.29 – 3.25 (m, 1H).
15	Methyl gallate	L, S	δ_H 7.15 (s, 2H), 3.86 (s, 3H).
16	(*E*)-*p*-Coumaric acid 4-*O*-*β*-D-glucopyranoside	L	δ_H 7.63 (d, *J* = 8.8 Hz, 2H), 7.57 (d, *J* = 16.0 Hz, 1H), 7.15 (d, *J* = 8.8 Hz, 2H), 6.44 (d, *J* = 16.0 Hz, 1H), 5.14 (d, *J* = 7.5 Hz, 1H), 3.92 (dd, *J* = 12.4, 2.2 Hz, 1H), 3.74 (dd, *J* = 12.4, 5.7 Hz, 1H), 3.62 (ddd, *J* = 9.0, 5.7, 2.2 Hz, 1H), 3.58 (d, *J* = 9.0, 1H), 3.55 (dd, *J* = 9.8, 7.5 Hz, 1H), 3.48 (dd, *J* = 9.8, 9.0 Hz, 1H).
20	Syringic acid *O*-*β*-D-glucopyranosyl ester (Erigeside C)	L, S	δ_H 7.46 (s, 2H), 5.76 (d, *J* = 7.8 Hz, 1H), 3.92 (s, 6H), 3.92 – 3.90 (m, 1H), 3.74 (m, 1H), 3.64 – 3.32 (overlapping)
21	Catechin	S	δ_H 6.92 (d, *J* = 2.1 Hz, 1H), 6.90 (d, *J* = 8.3 Hz, 1H), 6.84 (dd, *J* = 8.3, 2.1 Hz, 1H), 6.07 (d, *J* = 2.2 Hz, 1H), 6.00 (d, *J* = 2.2 Hz, 1H), 4.21 (td, *J* = 7.5, 5.4 Hz, 1H), 2.84 (dd, *J* = 16.1, 5.4 Hz, 1H), 2.54 (dd, *J* = 16.1, 7.8 Hz, 1H).

23	Taxifolin 7-*O*-*β*-D-glucopyranoside	S	δ_H 7.09 (d, *J* = 2.0 Hz, 1H), 7.00 (dd, *J* = 8.4, 2.0 Hz, 1H), 6.97 (d, *J* = 8.4 Hz, 1H), 6.32 (d, *J* = 2.2 Hz, 1H), 6.23 (d, *J* = 2.2 Hz, 1H), 5.16 (d, *J* = 11.9 Hz, 1H), 5.14 (d, *J* = 7.6 Hz, 1H), 4.82 (m, 1H).
26	Megastigman-7-ene-6,9,10-triol-3-one 9-*O*-*β*-D-glucopyranoside	L	δ_H 5.99 (d, *J* = 15.9 Hz, 1H), 5.71 (dd, *J* = 15.9, 7.9 Hz, 1H), 4.53 (m, 1H), 4.47 (d, *J* = 7.9 Hz), 3.68 – 3.72 (m, overlapping), 2.84 (d, *J* = 14.0 Hz, 1H), 2.47 (d, *J* = 13.4 Hz, 1H), 2.43 (m, 1H), 2.25 (dd, *J* = 13.4, 2.3 Hz, 1H), 1.95 (dd, *J* = 14.0, 2.3 Hz, 1H), 0.96 (s, 3H), 0.95 (d, *J* = 6.6 Hz, 3H), 0.93 (s, 3H).
28	Eucomic acid	S	δ_H 7.12 (d, *J* = 8.5 Hz, 2H), 6.80 (d, *J* = 8.5 Hz, 2H), 2.97 (d, *J* = 13.8 Hz, 1H), 2.93 (d, *J* = 16.1 Hz, 1H), 2.86 (d, *J* = 13.8 Hz, 1H), 2.66 (d, *J* = 16.1 Hz, 1H).
29	(*Z*)-*p*-Coumaric acid 4-*O*-*β*-D-glucopyranoside	L, S	δ_H 7.46 (d, *J* = 8.8 Hz, 2H), 7.08 (d, *J* = 8.8 Hz, 2H), 6.60 (d, *J* = 12.6 Hz, 1H), 5.98 (d, *J* = 12.6 Hz, 1H), 5.11 (d, *J* = 7.7 Hz, 1H), 3.91 (dd, *J* = 12.4, 2.2 Hz, 1H), 3.74 (dd, *J* = 12.4, 5.7 Hz, 1H), 3.61 (m, 1H), 3.58 (d, *J* = 9.0 Hz, 1H), 3.54 (dd, *J* = 9.3, 7.7 Hz, 1H), 3.48 (m, 1H).
30	Syringin	S	δ_H 6.86 (s, 2H), 6.59 (d, *J* = 15.8 Hz, 1H), 6.40 (dt, *J* = 15.8, 5.7 Hz, 1H), 4.98 (d, *J* = 7.5 Hz, 1H), 4.25 (dd, *J* = 5.7, 1.3 Hz, 2H), 3.87 (s, 6H).
33	Quercetin 3-*O*-rutinoside-7-*O*-*β*-D-glucopyranoside	L	δ_H 7.69 (d, *J* = 2.1 Hz, 1H), 7.64 (dd, *J* = 8.6, 2.1 Hz, 1H), 7.02 (d, *J* = 8.6 Hz, 1H), 6.86 (d, *J* = 2.1, 1H), 6.59 (d, *J* = 2.1 Hz, 1H), 5.23 (d, *J* = 7.6 Hz, 1H), 4.96 (d, *J* = 8.6 Hz, 1H), 4.54 (d, *J* = 1.3 Hz, 1H), 1.05 (d, *J* = 6.3 Hz, 3H).
34	3,4,5-Trimethoxyphenol 1-*O*-*β*-D-apiofuranosyl-(1→6)-*β*-D-glucopyranoside (Kelampayoside A)	S	δ_H 6.50 (s, 2H), 5.04 (d, *J* = 3.2 Hz, 1H), 5.03 (d, *J* = 7.8 Hz, 1H), 3.95 (d, *J* = 10.1 Hz, 1H), 3.92 (m, 1H), 3.85 (s, 6H), 3.81 (d, *J* = 10.1 Hz, 1H), 3.73 (s, 3H), 3.72 (m, 1H), 3.57 (s, 2H).
36	(*Z*)-Sinapic acid O-*β*-D-glucopyranoside	L	δ_H 7.34 (s, 2H), 6.95 (m, 1H), 5.98 (d, *J* = 12.6 Hz, 1H), 5.19 (d, *J* = 7.8 Hz, 1H), 3.90 (*s*, 6H).
38	(*Z*)-*p*-Coumaric acid 4-*O*-(2′-*O*-β-D-apiofuranosyl)-*β*-D-glucopyranoside	L, S	δ_H 7.48 (d, *J* = 8.8 Hz, 2H), 7.06 (d, *J* = 8.8 Hz, 2H), 6.70 (d, *J* = 12.6 Hz, 1H), 5.98 (d, *J* = 12.6 Hz, 1H), 5.40 (d, *J* = 2.3 Hz, 1H), 5.19 (d, *J* = 7.7 Hz, 1H), 4.00 (d, *J* = 2.3 Hz, 1H), 4.00 (d, *J* = 10.1 Hz, 1H), 3.85 (d, *J* = 10.1 Hz, 1H), 3.73 (dd, *J* = 12.5, 5.8 Hz, 1H), 3.65 (dd, *J* = 9.3, 7.7 Hz, 1H), 3.57 (s, 2H).
41	Quercetin 3-*O*-(6″-*O*-malonyl-*β*-D-glucopyranosyl)-7-*O*-*β*-D-glucopyranoside	L	δ_H 7.65 (d, *J* = 2.1 Hz, 1H), 7.59 (dd, *J* = 8.5, 2.1 Hz, 1H), 7.01 (d, *J* = 8.5 Hz, 1H), 6.88 (d, *J* = 2.1 Hz, 1H), 6.61 (d, *J* = 2.1 Hz, 1H), 5.24 (d, *J* = 7.3 Hz, 1H), 4.98 (d, *J* = 7.8 Hz, 1H), 4.16 (dd, *J* = 12.0, 2.0 Hz, 2H), 4.08 (dd, *J* = 12.0, 5.0 Hz, 1H), 3.81 (dd, *J* = 12.4, 2.0 Hz, 1H), 3.68 (dd, *J* = 12.4, 5.2 Hz, 1H).
44	Epigallocatechin gallate	L, S	δ_H 6.99 (s, 2H), 6.61 (s, 2H), 6.14 (d, *J* = 2.3 Hz, 1H), 6.10 (d, *J* = 2.3 Hz, 1H), 5.55 (m, 1H), 5.14 (m, 1H), 3.06 (dd, *J* = 17.2, 4.4 Hz, 1H), 2.91 (m, 1H).
47	Apigenin 6,8-di-*C*-*β*-D-glucopyranoside	L, S	Major rotamer (aglycone): δ_H 8.00 (d, *J* = 8.7 Hz, 2H), 7.04 (d, *J* = 8.7 Hz, 2H), 6.76 (s, 1H); Minor rotamer (aglycone): 7.92 (d, *J* = 8.6 Hz, 2H), 7.04 (d, *J* = 8.6 Hz, 2H), 6.76 (s, 1H); Glucoses: 5.20 (d, *J* = 9.9 Hz, 1H), 5.08 (d, *J* = 10.0 Hz, 1H), 5.08 (d, *J* = 10.0 Hz, 1H), 5.01 (d, *J* = 10.0 Hz, 1H), 3.97 – 3.55 (overlapping)
55	Apigenin (6-*C*-*α*-L-arabinopyranosyl)-8-*C*-*β*-D-glucopyranoside	L, S	Major rotamer (aglycone): δ_H 8.00 (d, *J* = 8.7 Hz, 2H), 7.04 (d, *J* = 8.7 Hz, 2H), 6.76 (s, 1H); Minor rotamer (aglycone): 7.92 (d, *J* = 8.6 Hz, 2H), 7.04 (*d*, *J* = 8.6 Hz, 2H), 6.76 (s, 1H); Sugars: 5.18 (d, *J* = 9.7 Hz, 1H), 5.09 (d, *J* = 10.0 Hz, 1H), 5.09 (d, *J* = 10.0 Hz, 1H), 4.95 (d, *J* = 9.7 Hz, 1H), 4.18 – 3.57 (overlapping)
56	Quercetin 3-*O*-di-*α*-L-rhamnopyranosyl-*β*-D-glucopyranoside	L, S	Major rotamer (aglycone): δ_H 7.76 (d, *J* = 2.1 Hz, 1H), 7.64 (dd, *J* = 8.4, 2.1 Hz, 1H), 7.05 (d, *J* = 8.4 Hz, 1H), 6.58 (d, *J* = 2.2 Hz, 1H), 6.35 (d, *J* = 2.2 Hz, 1H); Minor rotamer (aglycone): 7.74 (d, *J* = 2.1 Hz, 1H), 7.66 (dd, *J* = 8.6, 2.1 Hz, 1H), 7.03 (d, *J* = 8.6 Hz, 1H), 6.52 (d, *J* = 2.2 Hz, 1H), 6.35 (d, *J* = 2.2 Hz, 1H); Sugars: 5.31 (d, *J* = 7.5 Hz, 1H), 5.27 (d, *J* = 7.9 Hz, 1H), 5.12 (d, *J* = 1.2 Hz, 1H), 4.53 (d, *J* = 1.4 Hz, 1H), 4.19 – 3.23 (overlapping), 1.17 (d, *J* = 6.2 Hz, 3H), 1.08

			(d, *J* = 6.3 Hz, 3H), 1.05 (d, *J* = 6.3 Hz, 3H), 0.75 (d, *J* = 6.3 Hz, 3H).
58	Quercetin 3-*O*-*α*-L-rhamnopyranosyl-(1→2)-*β*-D-galactopyranoside	L, S	Major rotamer: δ_H 7.74 (d, *J* = 2.2 Hz, 1H), 7.63 (dd, *J* = 8.5, 2.2 Hz, 1H), 7.04 (d, *J* = 8.5 Hz, 1H), 6.58 (d, *J* = 2.1 Hz, 1H), 6.35 (d, *J* = 2.1 Hz, 1H), 5.45 (d, *J* = 7.7 Hz, 1H), 5.12 (d, *J* = 1.3 Hz, 1H), 0.99 (d, *J* = 6.4 Hz, 3H); Minor rotamer: 7.74 (d, *J* = 2.3 Hz, 1H), 7.67 (dd, *J* = 8.6, 2.2 Hz, 1H), 6.98 (d, *J* = 8.6 Hz, 1H), 6.53 (d, *J* = 2.2 Hz, 1H), 6.34 (d, *J* = 2.2 Hz, 1H), 5.03 (d, *J* = 7.4 Hz, 1H), 4.90 (d, *J* = 1.4 Hz, 1H), 0.88 (d, *J* = 6.3 Hz, 3H), 3.92 – 3.37 (overlapping)
59	Myricetin 3-*O*-(6″-*O*-malonyl)-*β*-D-glucopyranoside	L, S	δ_H 7.22 (s, 2H), 6.61 (d, *J* = 2.1 Hz, 1H), 6.36 (d, *J* = 2.1 Hz, 1H), 4.96 (d, *J* = 7.8 Hz, 1H), 4.18 (dd, *J* = 12.1, 2.0 Hz, 1H), 4.08 (dd, *J* = 12.1, 4.9 Hz, 1H), 3.55 (dd, *J* = 8.9, 7.9 Hz, 1H), 3.47 (m, 1H), 3.44 (m, 1H), 3.30 (m, 1H).
60	Quercetin 3-*O*-*α*-L-rhamnopyranosyl-(1→6)-*β*-D-galactopyranoside	L	Major rotamer: δ_H 7.69 (d, *J* = 2.1 Hz, 1H), 7.64 (dd, *J* = 8.4, 2.1 Hz, 1H), 7.03 (d, *J* = 8.4 Hz, 1H), 6.59 (d, *J* = 2.1 Hz, 2H), 6.35 (d, *J* = 2.1 Hz, 1H), 5.47 (d, *J* = 7.3 Hz, 1H), 5.14 (d, *J* = 1.7 Hz, 1H), 1.01 (d, *J* = 6.2 Hz, 3H); Minor rotamer: 7.74 (d, *J* = 2.2 Hz, 1H), 7.63 (dd, *J* = 8.5, 2.1 Hz, 1H), 7.04 (d, *J* = 8.5 Hz, 1H), 6.58 (d, *J* = 2.2 Hz, 1H), 6.35 (d, *J* = 2.2 Hz, 1H), 5.45 (d, *J* = 7.4 Hz, 1H), 5.11 (d, *J* = 1.8 Hz, 1H), 0.99 (d, *J* = 6.6 Hz, 3H).
61	Di-*O*-methyltaxifolin-*β*-D-glucopyranoside	S	δ_H 7.24 (d, *J* = 1.9 Hz, 1H), 7.17 (dd, *J* = 8.3, 1.9 Hz, 1H), 7.12 (d, *J* = 8.3 Hz, 1H), 6.33 (d, *J* = 2.2 Hz, 1H), 6.25 (d, *J* = 2.2 Hz, 1H), 5.23 (d, *J* = 11.9 Hz, 1H), 5.14 (d, *J* = 7.5 Hz, 1H), 4.87 (d, *J* = 11.9 Hz, 1H), 3.90 (s, 3H), 3.89 (s, 3H), 3.89 (dd, *J* = 12.3, 2.3 Hz, 1H), 3.73 (dd, *J* = 12.3, 5.5 Hz, 1H), 3.62 (ddd, *J* = 9.8, 5.4, 2.3 Hz, 1H), 3.59 – 3.52 (m, 2H), 3.48 (dd, *J* = 9.6, 9.0 Hz, 1H)
62	Taxifolin	S	δ_H 7.08 (d, *J* = 2.0 Hz, 1H), 6.99 (dd, *J* = 8.3, 2.0 Hz, 1H), 6.96 (d, *J* = 8.3 Hz, 1H), 6.08 (d, *J* = 2.1 Hz, 1H), 6.00 (d, *J* = 2.1 Hz, 1H), 5.10 (d, *J* = 11.7 Hz, 1H)
63	Quercetin 3-*O*-neohesperidoside	L, S	δ_H 7.76 (d, *J* = 2.2 Hz, 1H), 7.63 (dd, *J* = 8.5, 2.2 Hz, 1H), 7.02 (d, *J* = 8.5 Hz, 1H), 6.59 (d, *J* = 2.1 Hz, 1H), 6.36 (d, *J* = 2.1 Hz, 1H), 4.80 (m, overlapped), 4.52 (d, *J* = 1.5 Hz, 1H), 3.81 (m, 1H), 3.80 (dd, *J* = 10.0, 7.9 Hz, 1H), 3.68 (dd, *J* = 10.9, 3.3 Hz, 1H), 3.63 (m, 1H), 3.59 (dd, *J* = 9.9, 3.3 Hz, 1H), 3.55 – 3.41 (overlapping), 1.11 (d, *J* = 6.3 Hz, 3H).
65	Rutin	L, S	δ_H 7.66 (d, *J* = 2.2 Hz, 1H), 7.61 (dd, *J* = 8.5, 2.2 Hz, 1H), 7.01 (d, *J* = 8.5 Hz, 1H), 6.59 (d, *J* = 2.1 Hz, 1H), 6.36 (d, *J* = 2.1 Hz, 1H), 4.94 (d, *J* = 7.8 Hz, 1H), 4.54 (d, *J* = 1.5 Hz, 1H), 3.76 (dd, *J* = 11.1, 1.1 Hz, 1H), 3.66 (dd, *J* = 3.4, 1.7 Hz, 1H), 3.55 (dd, *J* = 9.4, 7.8 Hz, 1H), 3.50 (dd, *J* = 9.7, 3.4 Hz, 1H), 3.45 (t, *J* = 9.1 Hz, 1H), 3.43 – 3.37 (m), 3.35 (dd, *J* = 6.6, 1.5 Hz, 1H), 1.05 (d, *J* = 6.2 Hz, 3H).
66	Quercetin 3-*O*-*β*-D-galactopyranoside	L, S	δ_H 7.69 (d, *J* = 2.2 Hz, 1H), 7.60 (dd, *J* = 8.5, 2.2 Hz, 1H), 7.02 (d, *J* = 8.5 Hz, 1H), 6.58 (d, *J* = 2.2 Hz, 1H), 6.36 (d, *J* = 2.2 Hz, 1H), 4.90 (d, *J* = 7.8 Hz, 1H), 3.86 (d, *J* = 3.5 Hz, 1H), 3.77 (dd, *J* = 9.9, 7.8 Hz, 1H), 3.62 – 3.57 (m), 3.56 – 3.52 (m), 3.49 – 3.43 (m).
68	Epitaxifolin	S	δ_H 7.03 (d, *J* = 2.0 Hz, 1H), 6.93 (dd, *J* = 8.0, 2.0 Hz, 1H), 6.91 (d, *J* = 8.0 Hz, 1H), 6.10 (d, *J* = 2.1 Hz, 1H), 6.06 (d, *J* = 2.1 Hz, 1H), 5.49 (d, *J* = 3.1 Hz, 1H), 4.42 (d, *J* = 3.1 Hz, 1H).
70	Quercetin 3-*O*-*β*-D-glucopyranoside	L, S	δ_H 7.64 (d, *J* = 2.2 Hz, 1H), 7.59 (dd, *J* = 8.4, 2.2 Hz, 1H), 7.02 (d, *J* = 8.4 Hz, 1H), 6.59 (d, *J* = 2.1 Hz, 1H), 6.36 (d, *J* = 2.1 Hz, 1H), 4.99 (d, *J* = 7.8 Hz, 1H), 3.66 (dd, *J* = 12.3, 2.2 Hz, 1H), 3.54 (dd, *J* = 12.3, 5.2 Hz, 1H), 3.52 (dd, *J* = 9.3, 7.8 Hz, 1H), 3.44 (t, *J* = 9.3 Hz, 1H), 3.40 – 3.36 (m).
73	Quercetin 3-*O*-(6″-*O*-malonyl)-*β*-D-galactopyranoside	L, S	δ_H 7.68 (d, *J* = 2.2 Hz, 1H), 7.59 (dd, *J* = 8.5, 2.2 Hz, 1H), 7.01 (d, *J* = 8.5 Hz, 1H), 6.61 (d, *J* = 2.1 Hz, 1H), 6.37 (d, *J* = 2.1 Hz, 1H), 4.80 (m, 1H), 4.21 (dd, *J* = 11.6, 8.1 Hz, 1H), 4.03 (dd, *J* = 11.6, 4.1 Hz, 1H), 3.87 (d, *J* = 3.5 Hz, 1H), 3.78 (dd, *J* = 9.8, 7.8 Hz, 1H), 3.68 (m, 1H), 3.59 (dd, *J* = 9.9, 3.5 Hz, 1H).

74	Phloretin 3′,5′-di-*C*-*β*-D-glucopyranoside	S	δ_H 7.13 (d, *J* = 8.5 Hz, 2H), 6.79 (d, *J* = 8.5 Hz, 2H), 4.98 (d, *J* = 9.9 Hz, 2H), 3.92 – 3.88 (m, 2H), 3.86 – 3.80 (m, 2H), 3.70 (m, 2H), 3.62 – 3.57 (m), 2.92 (t, *J* = 7.6 Hz, 2H).
76	Quercetin 3-*O*-(6″-*O*-malonyl)-*β*-D-glucopyranoside	L, S	δ_H 7.59 (d, *J* = 2.1 Hz, 1H), 7.52 (dd, *J* = 8.5, 2.1 Hz, 1H), 6.96 (d, *J* = 8.5 Hz, 1H), 6.56 (d, *J* = 1.7 Hz, 1H), 6.34 (d, *J* = 1.7 Hz, 1H), 4.94 (d, *J* = 7.8 Hz, 1H), 4.16 (dd, *J* = 12.0, 2.0 Hz, 1H), 4.08 (dd, *J* = 12.0, 4.7 Hz, 1H), 3.87 (m 1H), 3.54 (dd, *J* = 8.9, 8.1 Hz, 1H), 3.48 – 3.36 (m, overlapping)
81	Kaempferol 3-*O*-*β*-D-glucopyranoside	L	δ_H 8.03 (d, *J* = 8.9 Hz, 2H), 7.01 (d, *J* = 8.9 Hz, 2H), 6.59 (d, *J* = 2.1 Hz, 1H), 6.36 (d, *J* = 2.1 Hz, 1H), 4.94 (d, *J* = 8.1 Hz, 1H).
88	Kaempferol 3-*O*-(6″-*O*-malonyl)-*β*-D-glucopyranoside	L	δ_H 7.99 (d, *J* = 8.9 Hz, 2H), 6.99 (d, *J* = 8.9 Hz, 2H), 6.60 (d, *J* = 2.1 Hz, 1H), 6.37 (d, *J* = 2.1 Hz, 1H), 4.93 (d, *J* = 7.7 Hz, 1H), 4.14 (dd, *J* = 12.1, 1.8 Hz, 1H), 4.09 (dd, *J* = 12.1, 4.4 Hz, 1H).
94	Di-*O*-methyltaxifolin	L, S	δ_H 7.23 (d, *J* = 1.9 Hz, 1H), 7.17 (dd, *J* = 8.2, 1.9 Hz, 1H), 7.12 (d, *J* = 8.2 Hz, 1H), 6.09 (d, *J* = 2.1 Hz, 1H), 6.02 (d, *J* = 2.1 Hz, 1H), 5.19 (d, *J* = 11.7 Hz, 1H), 3.90 (s, 3H), 3.89 (s, 3H).

Data collected at 600 MHz in D_2O:CD_3OD (4:1). Spectra were referenced to TSP-d_4 (0.01% w/v) at δ_H 0.00. Only clearly observed NMR signals are presented for each compound. Multiplicities and coupling constants (*J*) in Hz are given in parentheses. Extracts – L (leaves) and S (stems)

Annotation of other compounds

Generally, diverse classes of compounds were isolated and identified from both the leaf and stem extracts. Most of the compounds identified were flavonoids (constituting over half of the characterised compounds) and phenolic acid derivatives. Smaller numbers of other phenyl derivatives, organic acids, amino acids, alkyl glycosides, terpene derivatives and fatty acids were also identified.

Flavonoids:

The flavonoid composition of both extracts consisted mainly of glycosylated derivatives of various aglycones from multiple classes, with the majority being mono- or di- glycosylated compounds. Some tri-glycosylated derivatives as well as individual aglycones were also identified. Most compounds were flavonol derivatives mainly based on quercetin, but flavanol, flavone, flavanone and flavanonol derivatives were also identified.

Flavonol derivatives

The majority of the flavonol derivatives identified were quercetin glucosides. From the MS/MS fragmentation patterns, most of these were monodesmodic quercetin-3-*O*-glucosyl derivatives as shown by the presence of the base peak ion [M-2H-sugars]$^-$ at *m/z* 300 compared to the ion [M-H-sugars]$^-$ at *m/z* 301[28]. This was confirmed from the ^{1}H NMR resonances of both *meta*-coupled protons, H-6 and H-8, of most flavonoids in the region of 6.35 and 6.60 ppm compared to 6.60 and 6.90 ppm for the same protons in 7-*O*-glucosylated flavonoids. Some of the derivatives (**33** and **41**) included bidesmodic glucosides with substitution in both 3- and 7-positions of quercetin. Five malonylated glycosides of quercetin (**41**, **73**, **76**, **78** and **84**) were identified, as revealed by the neutral loss of *m/z* 44 followed by another of *m/z* 42 in their MS/MS spectra corresponding to a decarboxylation and loss of CH_2CO, respectively. These losses are characteristic for compounds bearing a malonyl group. However, the NMR signal of the malonyl methylene protons was not observed in the spectra of all the compounds due to an overlap with the residual methanol solvent signal as was similarly observed by Kazuma et al., (2003).[29] Thus, the presence of the malonyl moiety was assigned based only on the UHPLC-

MS data. Nevertheless, the attachment of the malonyl group in compounds **41**, **73** and **76** was confirmed to be on the 6-position of the sugar from the downfield shift in the chemical shifts of the protons at this position. For instance, a shift from δ_H 3.66/3.54 ppm in Quercetin 3-*O*-*β*-D-glucopyranoside (**70**) to δ_H 4.16/4.08 ppm in **76**. Quercetin 3-*O*-(6″-*O*-malonyl)-*β*-D-glucopyranoside (**76**) was suggested to be the major flavonoid in the leaf crude extract and appeared as the most abundant phenolic compound in the chromatogram of the extract. Compound **73**, also appeared in spectra from both the leaf and stem extracts and was similar to **76** but possessed a galactose instead of a glucose, with the same malonylation pattern. The compounds **78** and **84** could only be putatively identified as isomers of quercetin 3-*O*-malonylhexoside as their ^{1}H NMR data was not obtained. Quercetin derivatives have been reported from *Vachellia* and *Acacia* species,[10, 30-31] but as far as we know, no malonylated compounds based on quercetin or any other aglycone skeletons have been reported in any species belonging to the two genera before. Both the type of extract studied, and the methodology employed could be responsible for the annotation of malonylated flavonoids as encountered here when compared to common methodology and extracts of most reports in the literature. Quercetin and its derivatives are known to have anti-inflammatory activity and their abundance in both extracts could potentially explain the use of the plant in traditional medicine for relieving symptoms of cough, bronchitis, and measles.[9]

Compounds **67**, **79**, **85** and **89** all showed a base peak ion, [M-2H-sugars]$^-$ at *m/z* 314 (as compared to the ion [M-H-sugars]$^-$ at *m/z* 315) in their MS/MS indicating the presence of an isorhamnetin aglycone with a 3-*O*-glycosylation. Compound **85** was confirmed to be isorhamnetin 3-*O*-*β*-D-glucopyranoside by comparison with an authentic standard. The MS/MS spectrum of compound **79** showed an ion at *m/z* 315.0511 resulting from the loss of *m/z* 308 corresponding to a combined loss of a rhamnose and a hexose. The sugar sequence could not be defined as the compound showed unresolvable ^{1}H NMR data and a putative identification of isorhamnetin 3-*O*-rhamnosyl hexoside was given. Likewise, the mass spectrum of compound **67** showed its [M-H]$^-$ ion at *m/z* 769.2185, higher by 146 Da from that of compound **79** indicating the presence of an additional rhamnose in **67**. Without clear ^{1}H NMR signals, as in **79**, to confirm the sugar moieties, **67** was putatively identified as isorhamnetin 3-*O*-dirhamnosyl hexoside. Similarly, compound **89** showed a fragment at *m/z* 315.0510 corresponding to two successive neutral losses of *m/z* 86 and *m/z* 162 for malonyl and hexose moieties, respectively. This was hence putatively identified as isorhamnetin 3-*O*-malonyl hexoside, a compound that is described in *Vachellia* and *Acacia* species for the first time. The above mentioned isorhamnetin derivatives were identified only in the stem extract except for compound **89** which was identified in both the leaves and stems. Generally, isorhamnetin derivatives appear to have a rare occurrence in these genera according to the available studies on their phytochemical profiles.

Compounds **51** – **54** and **59** all showed a major fragment ion [M-2H-sugars]$^-$ at *m/z* 316 in their MS/MS indicating the presence of a myricetin aglycone that is 3-*O*-glycosylated. Compounds **52** and **53** showed the fragment ions [M-H-162]$^-$ and [M-H-308]$^-$ that are characteristic of flavonoid *O*-hexosides and *O*-rutinosides, respectively, and were indeed confirmed to be myricetin 3-*O*-*β*-D-glucopyranoside (**52**) and myricetin 3-*O*-rutinoside (**53**) by comparison with authentic standards. Compound **54** showed the same fragmentation pattern as **52** but with unresolvable ^{1}H NMR data, it was putatively identified as myricetin 3-*O*-hexoside. Compound **59** was confirmed to be myricetin 3-*O*-(6″-*O*-malonyl)-*β*-D-glucopyranoside by inspecting its

^{1}H NMR spectrum (Table 2). Compound **51** showed successive neutral losses of *m/z* 86 and *m/z* 162 and its molecular formula was consistent with the putative identification as a myricetin 3-*O*-malonyl hexoside. Two kaempferol derivatives, **81** and **88** were isolated from only the leaves and confirmed by ^{1}H NMR data (Table 2), with **88** being malonylated. These malonylated myricetin and kaempferol derivatives are also reported for the first time in *Vachellia* and *Acacia* species.

Flavanol derivatives

Seven flavanol (**18**, **19**, **21**, **22**, **35**, **44** and **48**) derivatives of catechins including their oligomers and esters with gallic acid were identified as within the extracts. Catechin (**21**), epicatechin (**35**) and procyanidin B3 (**18**) were confirmed by comparison of their profiles with authentic standards. The MS/MS pattern of compound **19** showed two major product ions at *m/z* 125.0251 ($C_6H_5O_3^-$) and 179.0355 ($C_9H_7O_4^-$) consistent with the retro Diels-Alder fragmentation of (epi)gallocatechin. Compound **44** showed gallic acid and (epi)gallocatechin fragments in its MS/MS pattern and was confirmed by ^{1}H NMR data (Table 2) as epigallocatechin gallate. Based on key fragments in the MS/MS data, including those at *m/z* 125.0247, 169.0149 ($C_7H_5O_5^-$) and 305.0662 ($C_{15}H_{13}O_7^-$), compound **22** with an $[M-H]^-$ ion at *m/z* 761.1347 and molecular formula $C_{37}H_{30}O_{18}$ was consistent with the putative identification of (epi)gallocatechin-(epi)gallocatechin gallate. Compound **48** was also putatively identified as a B-type proanthocyanidin based on its MS/MS pattern. Though not many were identified in this study, condensed tannins are some of the best-known compounds in *Acacias*.[10] Catechin, epicatechin and their oligomers were identified only in the stem extracts while their galloyl derivatives were identified in both the leaf and stem extracts. These flavanol derivatives are common constituents in *Vachellia* species and have been identified in different plant parts including the leaves of *V. tortilis*,[32] the leaves, bark, flowers and pods of *V. nilotica*,[33] and the leaves of *V. karroo* and *V. xanthophloea*.[30] These compounds are known to have antioxidant and anti-inflammatory activities.[9]

Flavone derivatives

Apigenin, luteolin and tricin constituted the core molecules of flavone derivatives found in both extracts. The aglycones, apigenin (**96**) and luteolin (**93**), and their respective 7-*O*-*β*-D-glucopyranosides, **86** and **72**, respectively, were identified by comparison of their ^{1}H NMR and UHPLC-MS data with that of authentic standards. The two aglycones and their 7-*O*-*β*-D-glucopyranosides have also been identified in *V. nilotica*.[33] Compounds **47**, **49**, **55** and **64** showed a major peak in the MS/MS consistent with the ion $[M-H-120]^-$ together with an additional ion peak corresponding to $[M-H-90]^-$ which are characteristic of flavonoid *C*-glycosides.[34] Compounds **49** and **64** were putatively identified as apigenin 6,8-di-*C*-pentosyl hexoside and apigenin *C*-hexoside respectively.[34] Analysis of the NMR spectra of the purified fractions confirmed the identity of **47** and **55** as apigenin 6,8-di-*C*-*β*-D-glucopyranoside and apigenin (6-*C*-*α*-L-arabinopyranosyl)-8-*C*-*β*-D-glucopyranoside (Table 2). The MS/MS of compound **91** showed the major fragment at *m/z* 329.0665 corresponding to the formula $C_{17}H_{13}O_7^-$ ($[M-H-162]^-$). Considering the additional fragments at *m/z* 476.0953 and 314.0433 which correspond to a loss of methyl groups from the $[M-H]^-$ and $[M-H-162]^-$ ions respectively, **91** was putatively identified as tricin hexoside, present in the stems only. Tricin derivatives are however rare in *Vachellia*s and *Acacia*s with only one known derivative, tricin 4′-*O*-*β*-(6′′-hydroxycinnamic)-glucoside having been identified in *V. nilotica*.[35]

Flavanone derivatives

Eriocitrin (**69**) from the leaves, and naringenin (**95**) from the stems, were identified by comparison with authentic standards. Additionally, compound **57** with molecular formula $C_{21}H_{22}O_{10}$ showed major fragment ions $[M-H-120]^-$ at *m/z* 313.0718 and $[M-H-90]^-$ at *m/z* 343.0824 in the MS/MS spectrum, a pattern that is characteristic of flavonoid *C*-glycosides. This was putatively identified as naringenin *C*-hexoside and was observed only in the stems.[36]

Flavanonol derivatives

Besides aromadendrin (**83**), all the other flavanonol derivatives identified were based on taxifolin including free and methylated aglycones, as well as glycosylated entities. Compounds **62** and **68** both with the $[M-H]^-$ ion at *m/z* 303.051 corresponding to a compound with the molecular formula $C_{15}H_{12}O_7$, were identified as taxifolin and epitaxifolin (Table 2) respectively. Compounds **23** and **50** had the same molecular ion and both showed a fragment at *m/z* 303.05 corresponding to the product ion $[M-H-162]^-$ indicating that they were both taxifolin hexosides. Compound **23** was confirmed to be taxifolin 7-*O*-*β*-D-glucopyranoside by analysis of its ^{1}H NMR data (Table 2). Compound **87** showed an $[M-H]^-$ ion at *m/z* 317.0667 corresponding to a compound with the molecular formula $C_{16}H_{14}O_7$ for a mono-methylated taxifolin. The compound was hence putatively identified as methyltaxifolin. It is worth noting that the MS/MS of the parent ion of **87** afforded the product ions at *m/z* 152.0121 ($C_7H_4O_4^-$) and 165.0572 ($C_9H_9O_3^-$) indicating methylation of the compound′s B-ring. Compound **61** showed a product ion $[M-H-162]^-$ at *m/z* 331.0822 as the major fragment in its MS/MS spectrum in line with a dimethylated taxifolin that has lost a hexose. The ^{1}H NMR data confirmed the identity of **61** to be di-*O*-methyltaxifolin-*O*-*β*-D-glucopyranoside (Table 2). All the taxifolin derivatives were identified in the stems except for di-*O*-methyltaxifolin (**94**) which also showed minor traces in the UHPLC-MS of the leaves. There are very limited reports of taxifolin in *Vachellia* species and no methylated taxifolin derivatives seem to have been reported in the genus.

Phenolic acid derivatives

The major fragment in the MS/MS spectra of compounds **16**, **24**, **29** and **38** was at *m/z* 163.04 ($C_9H_7O_3^-$) indicating the presence of a coumaric acid moiety in each of these compounds. The observed base peaks in compounds **16** and **29** corresponded to the loss of a hexose $[M-H-162]^-$ while their individual ^{1}H NMR spectra exhibited signals of (*E*)- and (*Z*)-isomers *p*-coumaric acid 4-*O*-*β*-D-glucopyranoside, respectively (Table 2). The MS/MS of the parent ion $[M-H]^-$ of compound **38** showed successive losses of a pentose (132 Da) and hexose (162 Da). The identity of the sugars as apiose and glucose as well as the configuration of the *p*-coumaric acid moiety were determined by inspecting the ^{1}H NMR spectrum, confirming the structure of **38** to be (*Z*)-*p*-coumaric acid 4-*O*-(2′-*O*-*β*-D-apiofuranosyl)-*β*-D-glucopyranoside (Table 2). Compound **24** had the same molecular formula and fragmentation pattern as **38** but the NMR of the fraction was dominated by signals of other co-eluting compounds and conclusive identification could not be achieved. Therefore, **24** was putatively identified as *p*-coumaric acid pentosyl hexoside, probably an isomer of **38**. The MS/MS of compound **40** showed two successive losses of pentose fragments (132 Da) followed by a hexose fragment (162 Da) to yield a coumaroyl fragment from the $[M-H]^-$ ion and was thus putatively identified as *p*-coumaric acid dipentosyl hexoside. Whereas **16** and **40** were only identified in the leaves and **24** only in the stems, **29** and **38** were identified in both the leaves and stems. Compound **36**

from the leaf extract showed a major fragment at *m/z* 223.0609 ($C_{11}H_{11}O_5^-$) from the loss of a hexose from the $[M-H]^-$ ion. This fragment was identified as sinapic acid by ^1H NMR inspection and **36** was confirmed to be (*Z*)-sinapic acid *O*-*β*-D-glucopyranoside (Table 2).

Compounds **7** and **71** were identified as gallic and salicylic acids, respectively, by comparing their profiles with those of authentic standards. Additionally, the gallic acid derivatives **6** and **15** were identified as *β*-glucogallin and methyl gallate respectively by ^1H NMR inspection (Table 2). Methyl gallate was implicated as the major compound responsible for the antiplasmodial activity of the leaf extracts of *V. xanthophloea*.[37] The major fragment in the MS/MS spectra of compounds **20** and **37** was $C_9H_9O_5^-$ (*m/z* 197.04) coming from syringic acid. From the ^1H NMR spectrum of the isolated fraction, compound **20** was confirmed to be erigeside C (Table 2) whereas there was insufficient information for the complete assignment of **37** and it could only be putatively identified as a syringic acid derivative. Compound **45** showed a major fragment of $C_8H_7O_4^-$ at *m/z* 167.0353 corresponding to the product ion $[M-H-162]^-$ and was putatively identified as a vanillic acid hexoside. Whereas **37** and **71** were identified from only the leaves and stems respectively, compounds **6**, **7**, **15** and **20** were identified in both the leaf and stem extracts.

Other phenyl derivatives

Compounds **30, 34** and **74** were identified from the stems only and were confirmed by ^1H NMR data inspection to be syringin, kelampayoside A and phloretin-3′,5′-di-*C*-*β*-glucopyranoside, respectively (Table 2). Compound **31** showed a loss of a hexose in its MS/MS spectrum to yield a fragment at *m/z* 165.0560 ($C_9H_9O_3^-$). Its NMR was not clear enough but showed signals at δ_H 7.22 (d, J = 8.4 Hz, 2H) and 6.79 (d, J = 8.4 Hz, 2H) pointing to the presence of an aromatic core, and an anomeric signal most likely from the hexose at δ_H 5.05 (d, J = 7.6 Hz, 1H). This was hence putatively identified as a phenyl hexoside derivative. Compound **82** appearing in only the stems was putatively identified as the lignan derivative, pinoresinol hexoside by comparing its MS/MS pattern with that reported in literature.[38]

Other compounds

Two cyanogenic derivatives (**9** and **42**) were identified. The MS/MS of the $[M-H]^-$ ion at *m/z* 428.1193 of compound **42**, only present in the leaf extract, showed product ions at *m/z* 401.1093 ($C_{17}H_{21}O_{11}^-$) for a loss of HCN and at *m/z* 313.0571 ($C_{13}H_{13}O_9^-$) for the loss of 2,3-dihydroxy-3-methylbutanenitrile as in **9** alongside an additional loss of a galloyl moiety shown by the fragment at *m/z* 151.0044 ($C_7H_3O_4^-$). The presence of the galloyl moiety was confirmed by the appearance of a signal at δ_H 7.21 (s, 2H) in the ^1H NMR spectrum. Similarly signals for the 3-hydroxy-3-methylbutanenitrile moiety were present at δ_H 5.66 (s, 1H), 1.56 (s, 3H) and 1.52 (s, 3H). However, the ^1H NMR was not clear enough for conclusive identification and **42** was putatively identified as a cyanogenic derivative. The closest structure with similar characteristics found in the literature is linamarin gallate isolated from the Nigerian mistletoe *Loranthus micranthus* (Linn.).[39] Regardless of the position of the galloyl substituent on the core structure of the nitrile glucoside as shared by both compounds, **42** is potentially a new compound.

An additional alkyl alcohol glucoside (**8**) with a novel structure in nature was detected in the leaves showing the formate adduct, $[M+HCOOH-H]^-$, at *m/z* 297.1191 corresponding to a compound with the molecular formula $C_{10}H_{20}O_7$. Its ^1H NMR exhibited resonances typical of

a *β*-glucose in addition to aliphatic singlets at δ_H 1.39 (s, 3H) and 1.40 (s, 3H) among other signals. While highlighting the novelty of this compound, additional physical and chemical data is needed for its complete characterisation. Another metabolite (**10**) whose formate adduct at *m/z* 295.1034 corresponds to a compound with molecular formula $C_{10}H_{18}O_7$ was also isolated. This compound shared some similar MS/MS fragments with **9** including the major fragment at *m/z* 161.046 ($C_6H_9O_5^-$) pointing to a closeness in their structures. With no exploitable ^{1}H NMR data, **10** was putatively identified as an alkyl glucoside with its structure probably being similar to *β*-D-glucopyranosyl-2-methylpropanoate isolated from the flowers of *Moricandia arvensis*[40] or related analogues.

Amino-, fatty- and other organic acids were also identified in the extracts. Tryptophan (**14**) was identified by comparison with an authentic standard as were malic (**3**), citric (**4**) and succinic acids (**5**). Hydroxycitric acid (**1**), hibiscus acid (**2**), and the fatty acids, **97** – **102**, were identified putatively based on UHPLC-MS data. The presence of the only abscisic-like terpene, megastigman-7-ene-6,9,10-triol-3-one 9-*O*-*β*-D-glucopyranoside (**26**) was also confirmed in the leaf extract and not in the stems by ^{1}H NMR data comparison with a similar compound in the literature (Table 2).[41]

Conclusion

To the best of our knowledge, this paper provides the first characterisation of the chemical profile of *V. gummifera*, a plant with importance in traditional medicine. Over 100 metabolites have been identified and their occurrence in the leaves and/or stems has been reported. Our methodology led to the characterisation of a vast number of metabolite classes. Several compounds were present in both the leaf and stem extracts albeit differing in their abundances in the two tissues. However, eucomic acid and piscidic acid were some of the main compounds differentiating the leaf and stem profiles, being identified in only the latter. Pinitol, a cyclitol known to be an anti-diabetic agent, was the major compound in both leaves and stems. Other metabolites classes included amino, organic, and phenolic acids, in addition to flavonoid derivatives of flavonols, flavanols, flavones, flavanones and flavanonols. Many compounds described in our study are known plant natural products with valuable biological potential. The structures of the two potentially novel compounds that were isolated could not be confirmed and there is thus need for more elaborate analysis to fully characterise these structures. Additional studies are also necessary to provide an even more in-depth fingerprint of the plant′s metabolome for instance the less polar and non-polar compound profiles as well as the chemical profile of other plant parts such as the fruits (pods and seeds). Furthermore, studies are needed to investigate the bioactivity of *V. gummifera* extracts or its prominent compounds, and to ascertain which compounds contribute to the reported vernacular uses of the plant.

Experimental

Plant material and extraction

Leaves and whole stems (including the bark) of *V. gummifera* were harvested from two-year-old plants at the experimental farm (32.219731E, -7.892268N) of Mohammed VI Polytechnic University in Ben Guerir, Morocco in September 2019. The samples were air-dried in the dark

at room temperature for 20 days before being transported to Rothamsted Research in the UK where they were milled into a fine powder (Retsch Ultra Mill ZM200, Retsch, UK). The milled samples were stored at room temperature in the dark until analysis. Voucher specimens of the plant are available at the Mohammed VI Polytechnic University experimental farm.

Extraction was done using the method reported by Noleto-Dias et al., (2020)[42] with minor modifications. Briefly, for initial metabolite screening, triplicate replicates (15 mg) of each sample were suspended in either H_2O/CH_3OH (4:1 v/v, 1 mL) for UHPLC-MS or D_2O/CD_3OD (4:1 v/v) containing 3-(trimethylsilyl) propionic acid-d_4, 0.01 % *w*/*v* (TSP-d_4) (1 mL) (NMR solvent) for ^{1}H NMR. The samples were vortexed for 10 s and then heated at 50°C for 10 min. They were centrifuged at 13200 rpm for a further 10 min. The supernatants were transferred to clean tubes and heated at 90°C for 2 min. After, they were cooled at 4°C for 30 min and then centrifuged at 13200 rpm for 10 min. The supernatants were then transferred to glass vials for UHPLC-MS or 5 mm NMR tubes for NMR analysis. For fractionation, 260 mg of each sample was extracted in 6 mL of H_2O:CH_3OH (4:1 v/v) using the same procedure and the resultant extract was aliquoted into a glass autosampler vial for HPLC fractionation.

Fractionation

Fractionation was carried out using a Dionex UltiMate 3000, Thermo Fisher Scientific HPLC system equipped with an Ascentis C-18 column (5 μm, 5 × 250 mm, Supelco, UK). Chromatographic separation was performed using a constant flow rate of 1 mL/min of the mobile phases, water + 0.1% formic acid (A) and acetonitrile + 0.1% formic acid (B). The binary gradient was: 0–10 min, 5% B; 10–50 min, 22% B; 50–60 min, 37% B; 60–70 min, 50% B; 70–80 min, 70% B; and finally, 80–95 min, 100% B. Multiple injections (each 100 μL) were made, and the resultant fractions were automatically collected by time into individual glass tubes. The system was automatically set to restart the collection into the same glass tubes at each run. The eluting compounds were monitored between the wavelengths, 200 and 800 nm. At the end of the collection, each fraction (200 μL) was transferred into a glass vial and subsequently analysed by UHPLC-MS. The remaining volume was dried overnight using a Speedvac concentrator (Genevac, Suffolk, UK) and then dissolved in 700 μL of NMR solvent for subsequent NMR analysis.

UHPLC-MS analysis

UHPLC–MS data were recorded on an LTQ-Orbitrap Elite mass spectrometer coupled to a Dionex UltiMate 3000 RS UHPLC system (Fisher Scientific). Samples (10 μL) were injected onto a reversed-phase Hypersil GOLD C18 selectivity HPLC column (3 μm, 30 ×2.1 mm i.d. Thermo Fisher Scientific) maintained at 35 °C. The solvent system consisted of water/0.1% formic acid (A) and acetonitrile/0.1% formic acid (B). Total run time was 40 min using a flow rate of 0.3 mL/min and the following elution gradient: 0–5 min, 0% B; 5–27 min, 31.6% B; 27–34 min, 45% B; 34–37.5 min,75% B. Mass spectra were collected using a heated ESI source and mass spectra were acquired in negative mode with a resolution of 120,000 over *m/z* 50–1500. The source voltage, sheath gas, auxiliary gas, sweep gas and capillary temperature were set to 2.5 kV, 35 (arbitrary units), 10 (arbitrary units), 0.0 (arbitrary units) and 350 °C, respectively. Default values were used for other acquisition parameters. Automatic MS/MS fragmentation was performed on top four ions using an isolation width of *m/z* 2. Ions were fragmented using high-energy C-trap dissociation (HCD) with a normalised collision energy

of 65 and an activation time of 0.1 ms. Data was collected and inspected using Xcalibur v. 2.2 (Thermo Fisher Scientific).

NMR spectroscopy analysis

^{1}H NMR spectra were collected using a Bruker Avance 600 MHz NMR spectrometer (Bruker Biospin, Germany) operating at 600.05 MHz (^{1}H). Spectra were acquired at 300 K using a 5 mm TCI cryoprobe by using the zgpr pulse sequence with a 90°angle. Residual water was suppressed by pre-saturation during a 5 s delay. Spectra consisted of 64,000 data points and a spectral width of 12 ppm. FIDs were automatically transformed within Topspin version 4.2.0 (exponential window and a line broadening of 0.5 Hz). Phasing and baseline correction were carried out within the instrument software and chemical shifts were referenced relative to TSP-d_4. Where necessary, two-dimensional ^{1}H-^{1}H COSY, TOCSY and NOESY as well as ^{1}H-^{13}C HSQC and HMBC spectra were acquired to discriminate isomers or elucidate positions of substituents for certain compounds. Data was analysed using MestreNova software. Quantification of selected compounds was achieved via integration of characteristic multiplets in the ^{1}H NMR spectra.

Statistical analysis

The quantities of selected compounds were determined in triplicate. A *t*-test was used to assess whether the mean quantities of the compounds in the leaves and stems were statistically different from each other. Differences at $p < 0.05$ were considered significant. Analysis was done using Genstat (22nd edition, VSN International Ltd., Hemel Hempstead, U.K.).

Author Contributions

MK carried out the experimental work; MK, CN and GTMB analysed the data; IN collected and processed the plant material; MK wrote the manuscript; CN, GTMB, MA, LAT, MS, MHB and JLW reviewed the manuscript; all authors have read and approved the published version of the manuscript.

Acknowledgements

This research paper is part of the PhD project “Agricultural Bioproducts from Medicinal and Aromatic Plants” and sits within the FP05 project (Bioproducts for African Agriculture), which is funded by OCP Morocco. FP05 is a collaboration between Mohammed VI Polytechnic University, Rothamsted Research and Cranfield University. The authors would also like to acknowledge Charlotte Lomax and Gianluca Ruvo for their assistance in the collection of UHPLC-MS and NMR data respectively.

REFERENCES

[1] B. Kyalangalilwa, J. S. Boatwright, B. H. Daru, O. Maurin, M. van der Bank, *Bot. J. Linn. Soc.* **2013**, *172*, 500-523.

[2] Royal Botanic Gardens, Kew, "*Vachellia gummifera* (Willd.) Kyal. & Boatwr.", can be found under https://powo.science.kew.org/taxon/urn:lsid:ipni.org:names:77131720-1 (Accessed 05 May 2023).

[3] S. Belhaj, J. Dahmani, N. Belahbib, L. Zidane, *Ethnobot. Res. Appl.* **2020**, *20*, 1-40.
[4] F. Mouhajir, J. B. Hudson, M. Rejdali, G. H. N. Towers, *Pharm. Biol.* **2001**, *39*, 364-374.
[5] S. Belhaj, N. Chaachouay, L. Zidane, *J. Pharm. Pharmacogn. Res.* **2021**, *9*, 619-662.
[6] N. E. Allagui, S. Tahrouch, M. Bourijate, A. Hatimi, *Acta Bot. Gallica* **2007**, *154*, 503-509.
[7] M. A. Ait Baamrane, W. Shehzad, A. Ouhammou, A. Abbad, M. Naimi, E. Coissac, P. Taberlet, M. Znari, *PLoS ONE* **2012**, *7*, e35643.
[8] B. Jamoussi, R. Chakroun, C. Jablaoui, L. Rhazi, *Arabian J. Chem.* **2020**, *13*, 7459-7481.
[9] D. Jæger, M. C. O'Leary, P. Weinstein, B. L. Møller, S. J. Semple, *Phytochem. Rev.* **2019**, *18*, 129-172.
[10] D. S. Seigler, *Biochem. Syst. Ecol.* **2003**, *31*, 845-873.
[11] F. Mouhajir, J. A. Pedersen, M. Rejdali, G. H. N. Towers, *Pharm. Biol* **2001**, *39*, 391-398.
[12] M. Kisiriko, M. Anastasiadi, L. A. Terry, A. Yasri, M. H. Beale, J. L. Ward, in *Molecules* **2021**, *26*.
[13] R. Chaubal, P. V. Pawar, G. D. Hebbalkar, V. B. Tungikar, V. G. Puranik, V. H. Deshpande, N. R. Deshpande, *Chem. Biodiversity* **2005**, *2*, 684-688.
[14] E. Hernández-García, A. García, E. Garza-González, F. G. Avalos-Alanís, V. M. Rivas-Galindo, J. Rodríguez-Rodríguez, V. M. Alcantar-Rosales, C. Delgadillo-Puga, M. del Rayo Camacho-Corona, *J. Ethnopharmacol.* **2019**, *230*, 74-80.
[15] A. M. Kayed, E. A. M. Genady, H. A. Kadry, E.-S. M. Elghaly, *Nat. Prod. Res.* **2021**, *35*, 5571-5580.
[16] A. Pandi, V. M. Kalappan, N. Chandrashekar, *Bull. Natl. Res. Cent.* **2022**, *46*, 130.
[17] A. Pandi, K. Sattu, V. M. Kalappan, V. Lal, S. R. Varikasuvu, A. Ganguly, J. Prasad, *J. Food Biochem.* **2022**, *46*, e14282.
[18] A. Kumari, P. Das, A. K. Parida, P. K. Agarwal, *Front. Plant Sci.* **2015**, *6*.
[19] A. J. Borim de Souza, F. M. Ocampos, R. Catoia Pulgrossi, A. L. Dokkedal, L. A. Colnago, I. Cechin, L. L. Saldanha, in *Metabolites* **2023**,*13*.
[20] L. Brimer, S. B. Christensen, J. W. Jaroszewski, F. Nartey, *Phytochemistry* **1981**, *20*, 2221-2223.
[21] M. Okada, S. Park, T. Koshizawa, M. Ueda, *Tetrahedron* **2009**, *65*, 2136-2141.
[22] M. Rykaczewski, M. Krauze-Baranowska, J. Żuchowski, M. Krychowiak-Maśnicka, J. Fikowicz-Krośko, A. Królicka, *Phytochem. Lett.* **2019**, *30*, 245-253.
[23] W. Rosa, O. da Silva Domingos, P. P. de Oliveira Salem, I. S. Caldas, M. Murgu, J. H. G. Lago, P. Sartorelli, D. F. Dias, D. A. Chagas-Paula, M. G. Soares, *Phytochem. Anal.* **2021**, *32*, 859-883.
[24] M. Takahira, A. Kusano, M. Shibano, G. Kusano, T. Miyase, *Phytochemistry* **1998**, *49*, 2115-2119.
[25] G. T. M. Bitchagno, V.-A. Nchiozem-Ngnitedem, D. Melchert, S. A. Fobofou, *Nat. Rev. Chem.* **2022**, *6*, 806-822.
[26] N. Eiki, T. G. Manyelo, Z. M. Hassan, S. L. Lebelo, N. A. Sebola, B. Sakong, M. Mabelebele, *Sci. Rep.* **2022**, *12*, 21335.
[27] O. Purev, C. Purevsuren, S. Narantuya, S. Lkhagvasuren, H. Mizukami, A. Nagatsu, *Chem. Pharm. Bull.* **2002**, *50*, 1367-1369.
[28] C. Zhou, Y. Luo, Z. Lei, G. Wei, *Evid. Based Complement. Alternat. Med.* **2018**, *2018*, 8936307.
[29] K. Kazuma, N. Noda, M. Suzuki, *Phytochemistry* **2003**, *62*, 229-237.
[30] G. K. More, S. Meddows-Taylor, G. Prinsloo, *Molecules* **2021**, *26*.

[31] L. J. Rather, I. Shahid ul, F. Mohammad, *Sustainable Chem. Pharm.* **2015**, *2*, 12-30.
[32] B. E. C. Ziani, M. Carocho, R. M. V. Abreu, K. Bachari, M. J. Alves, R. C. Calhelha, O. Talhi, L. Barros, I. C. F. R. Ferreira, *Food Biosci.* **2020**, *36*, 100616.
[33] D. Zheleva-Dimitrova, K. I. Sinan, O. K. Etienne, G. Ak, J. B. Sharmeen, G. Dervisoglu, F. A. Ozdemir, M. F. Mahomoodally, G. Zengin, *Food Chem. Toxicol.* **2021**, *156*, 112446.
[34] P. Geng, J. Sun, M. Zhang, X. Li, J. M. Harnly, P. Chen, *J. Mass Spectrom.* **2016**, *51*, 914-930.
[35] Z. Khanam, F. Adam, O. Singh, J. Ahmad, *BioResources* **2011**, *6*, 2932-2940.
[36] S.-D. Chen, C.-J. Lu, R.-Z. Zhao, *Molecules* **2015**, *20*, 1594-1609.
[37] N. Tajuddeen, T. Swart, H. C. Hoppe, F. R. van Heerden, *Pharmaceuticals* **2022**,*15*.
[38] C. Li, C.-Z. Sun, Y.-H. Yang, N. Ma, Y.-J. Wang, F.-X. Zhang, Y.-H. Pei, *Phytomedicine* **2022**, *96*, 153888.
[39] M. O. Agbo, D. Lai, F. B. C. Okoye, P. O. Osadebe, P. Proksch, *Fitoterapia* **2013**, *86*, 78-83.
[40] H. Braham, H. Ben Jannet, Z. Mighri, P. M. Abreu, *J. Soc. Chim. Tunis.* **2007**, *9*, 109-113.
[41] H. Morikawa, R. Kasai, H. Otsuka, E. Hirata, T. Shinzato, M. Aramoto, Y. Takeda, *Chem. Pharm. Bull.* **2004**, *52*, 1086-1090.
[42] C. Noleto-Dias, C. Harflett, M. H. Beale, J. L. Ward, *Phytochem. Lett.* **2020**, *35*, 88-93.

Graphical Abstract