Rothamsted Research
Harpenden, Herts, AL5 2JQ

ROTHAMSTED RESEARCH

Telephone: +44 (0)1582 763133
Web: http://www.rothamsted.ac.uk/

Rothamsted Repository Download

A - Papers appearing in refereed journals

Kisiriko, M., Bitchagno-Mbahbou, G., Noleto-Dias, C., Naboulsi, I., Anastasiadi, M., Terry, L. A., Sobeh, M., Beale, M. H. and Ward, J. L. 2025. The untargeted metabolomic analysis of Ammodaucus leucotrichus Coss. & Dur. seeds reveals previously undescribed polar lignans and terpenoids. *Phytochemistry Letters.* 65, pp. 28-37. https://doi.org/10.1093/plphys/kiae584

The publisher's version can be accessed at:

- https://doi.org/10.1093/plphys/kiae584

The output can be accessed at: https://repository.rothamsted.ac.uk/item/992x2/the-untargeted-metabolomic-analysis-of-ammodaucus-leucotrichus-coss-dur-seeds-reveals-previously-undescribed-polar-lignans-and-terpenoids.

© 2025. This manuscript version is made available under the CC-BY-NC-ND 4.0 license http://creativecommons.org/licenses/by-nc-nd/4.0/

06/01/2025 12:20 repository.rothamsted.ac.uk library@rothamsted.ac.uk

Rothamsted Research is a Company Limited by Guarantee
Registered Office: as above. Registered in England No. 2393175.
Registered Charity No. 802038. VAT No. 197 4201 51.
Founded in 1843 by John Bennet Lawes.

Phytochemistry Letters 65 (2025) 28–37

Contents lists available at ScienceDirect

Phytochemistry Letters

journal homepage: www.elsevier.com/locate/phytol

The untargeted metabolomic analysis of *Ammodaucus leucotrichus* Coss. & Dur. seeds reveals previously undescribed polar lignans and terpenoids

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

ARTICLE INFO

Keywords:
Ammodaucus leucotrichus
Apiaceae
UHPLC-MS
NMR
Lignan
Terpene
Ammolactone-A
Ammodauside
Ammodaucic acid
Ammodolide

ABSTRACT

The polar metabolome of the medicinal plant *Ammodaucus leucotrichus* Coss. & Dur. has not been comprehensively characterised. In this study, the chemical composition of a water/methanol (4:1) extract of seeds was determined through a combination of UHPLC-MS and NMR techniques. Sixty compounds were identified from the extract with 36 of these confirmed using UHPLC-MS and/or NMR data while the remaining 24 were given putative identifications based on UHPLC-MS data. The compounds included lignans, terpenes, phenolics, flavonoids, and alkaloid derivatives together with some amino- and organic acids. Of these 2 terpenoids and 3 lignans were found to be novel and were isolated and structures determined by comprehensive NMR studies. A further novel lignan glycoside was tentatively identified. Together these data represent the most comprehensive profile of this traditional medicinal plant, providing an annotated profile that can form the basis of future correlative metabolomic investigations to determine active principles behind its reported bioactivities.

1. Introduction

Ammodaucus leucotrichus Coss. & Dur. (Apiaceae) is a flowering plant endemic to the North African Sahara region with presence in Morocco, Algeria, Tunisia, and Egypt, but also reported to extend to some sub-Saharan tropical regions (Idm'hand et al., 2020; Selama et al., 2022). In Morocco, it is locally known as "Kammûn es-sofi or akâman" where its different parts, especially the fruits and seeds, are used in traditional medicine mainly as infusions or decoctions to treat a wide range of ailments, including stomach and digestive diseases, diabetes, cough, cold, allergies and heart complications (Idm'hand et al., 2020). Several studies have reported additional biological effects from the seeds and aerial parts including *in vitro* antioxidant and antiglycation properties (Selama et al., 2022), *in vivo* antidiabetic, antihyperglycemic, and anti-inflammatory effects in mice (Es-Safi et al., 2020) as well as cytotoxic and antibacterial effects (Ziani et al., 2019). The hepatoprotective potential (Annaz et al., 2024) and *in vitro* antiproliferative properties (Belbachir et al., 2024) have also been reported.

Most of the chemical studies of this plant have focussed on the analysis of volatile compounds in essential oils and have commonly identified perillaldehyde and limonene as the two major volatiles in the oils of its fruits and seeds (Dahmane et al., 2017; Manssouri et al., 2020; Neghliz-Benabdelkader et al., 2021, Annaz et al., 2024). There are a few studies in the literature that have attempted to characterise the more polar metabolome of *A. leucotrichus*. Most of these have either only performed basic phytochemical screens, such as antioxidant activities and mineral analysis or have been largely incomplete, identifying only a few metabolites matching to available standards by HPLC (Bouknana et al., 2022; Es-Safi et al., 2020; Mouderas et al., 2020; Bouzaid et al., 2024a). A deeper study on the phenolic profile of the aerial parts using HPLC-DAD-ESI/MS^n reported a handful of flavonoids based on apigenin and luteolin, along with other phenolic compounds (Ziani et al., 2019). Subsequently, an untargeted UHPLC-MS/MS metabolomic analysis of the water and ethanol extracts of the plant's fruits reported 94 compounds but these could only be tentatively identified (Abderrezag et al., 2021). Even the most recent studies have employed variations of HPLC-MS methodologies to only putatively identify compounds, including phenolic acids and flavonoids (Abderrezag et al., 2024; Belbachir et al., 2024; Bouknana et al., 2024; Bouzaid et al., 2024b). A more targeted approach using diethyl ether extraction and GC-MS and NMR,

* Corresponding author.
E-mail address: jane.ward@rothamsted.ac.uk (J.L. Ward).

https://doi.org/10.1016/j.phytol.2024.11.006
Received 23 July 2024; Received in revised form 4 November 2024; Accepted 4 November 2024
Available online 21 November 2024
1874-3900/© 2024 Published by Elsevier Ltd on behalf of Phytochemical Society of Europe.

led to identification of a guaianolide, ammolactone-A together with some monoterpenoids including (+)-3-hydroxyperillaldehyde (Muckensturm et al., 1997). A further study has provided more complete spectroscopic characterisation and quantification of levels of ammolactone-A and *R*-perillaldehyde (Bellau et al., 2022) whilst more recently, polyacetylenic caffeoyl amides, in addition to some known flavones and lignans, were identified from the ethyl acetate extract of the fruits of *A. leucotrichus* using NMR and mass spectrometry (Hajib et al., 2023).

Given the extensive use of fruits/seeds of this plant to prepare medicinal teas and similar decoctions, there is a clear need to understand the chemical composition of more polar extractables. In this study we apply modern metabolomic technologies to reveal the full complexity of the chemical fingerprint that would be expected in such extracts. Structural data on 59 metabolites from seeds of *A. leucotrichus* using a combination of high-resolution UHPLC-MS and NMR, as well as the identification of 5 new metabolites, provides a much-needed detailed metabolite profile that can be used as a basis for future biomedical studies, including standardisation of herbal products derived from seeds of this plant.

2. Results and discussion

After initial collection of NMR and UHPLC-MS profiles of total extracts (water/methanol, 4:1) from seeds of *A. leucotrichus*, the crude extract was then split into 110 fractions by preparative reverse-phase HPLC. The crude fractions were each analysed in detail by UHPLC-MS and NMR. Known compounds were identified by comparing these data, including LC retention times, to our in-house spectral databases and/or to the literature (Table S1 and List S1 under Supplementary materials). For unknown metabolites, putative identifications were based on empirical formulae from high-resolution UHPLC-MS and MS/MS fragmentation patterns. Structures were then determined by combined analysis of these data, supplemented with data from both 1D and 2D-NMR experiments. Fifty-nine compounds were identified in UHPLC-MS/NMR data and one additional compound was identified based only

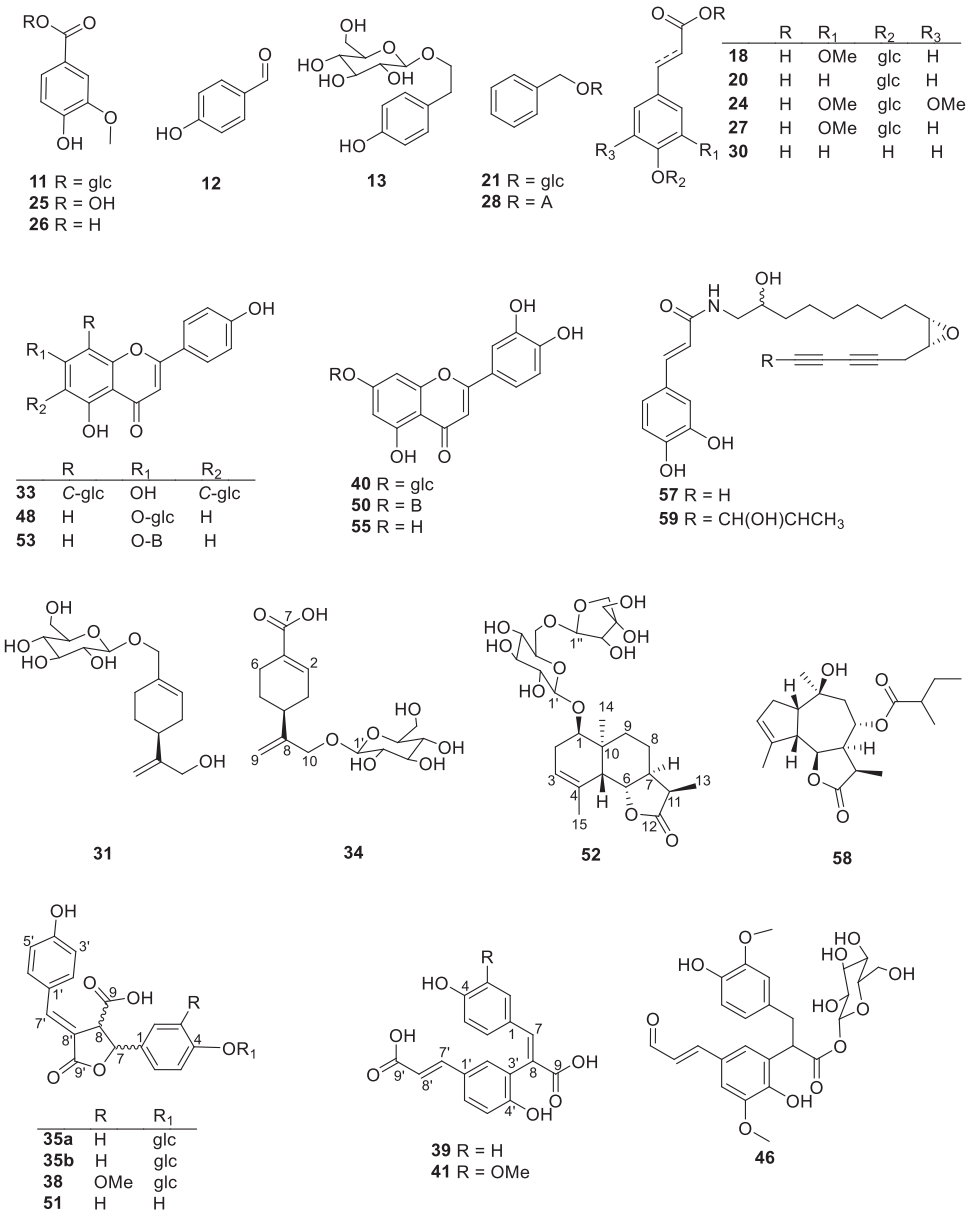

Fig. 1. Chemical structures of selected compounds isolated from seeds of *A. leucotrichus*. Abbreviations: glc = *β*-glucopyranose, Me = methyl, A = *β*-apiofuranosyl-(1→6)-*O*-*β*-glucopyranose, B = (6″-*O*-malonyl)-*β*-glucopyranose.

on its NMR data because it did not ionise in the LCMS system. The structures of selected compounds, including the novel metabolites, are shown in Fig. 1, and the UHPLC-MS profile of the crude extract is shown in Fig. 2. The UHPLC–MS/MS data with the level of identification of each compound (**1–59**) is summarised in Table S1.

Compound **34** was assigned a molecular formula of $C_{16}H_{24}O_8$ based on its deprotonated ion at *m/z* 343.1396 $[M\text{-}H]^-$ (calcd. for $C_{16}H_{23}O_8^-$, 343.1398). Its MS/MS pattern revealed the ion, $[M\text{-}H\text{-}162]^-$, at *m/z* 181.0877 ($C_{10}H_{13}O_3^-$) indicating the presence of a hexose. The ^{1}H NMR data of the compound, especially in the upfield region (Table 1), was similar to that of limonene-7,10-diol-7-*O*-*β*-glucopyranoside (komarovin B), a compound that was previously isolated from *Dracocephalum komarovi* (Toshmatov et al., 2019) and is reported in *A. leucotrichus* for the first time in this study (peak **31,** Fig. 2/Table 1/Table S1). The major differences between the NMR data of the two compounds (**31** and **34**) relate to the 7- and 10-positions (Table 1). The geminal hydrogens in the 10-position of **34** are deshielded compared to **31** and appear as doublets δ_H 4.43 (1H, d, J = 12.7 Hz, H-10a) and 4.24 (1H, d, J = 12.7 Hz, H-10b). Both H-10a and H-10b showed correlations to the anomeric carbon at δ_C 104.2 (C-1′) in the HMBC spectrum (Fig. 3), indicating the attachment of the sugar to the 10-position of the aglycone. The pattern of sugar signals and the anomeric proton's coupling constant (J = 8.3 Hz) indicated that the sugar was a *β*-glucopyranose. The assignment of C-8 at 151.6 was based on its HMBC correlations with both H-10a and H-10b as well as H-9a. The assignment of C-7 as a quaternary carbon was based on the absence of signals in the ^{1}H NMR spectrum that could be attributed to this position as in compound **31**. However, the MS/MS of the parent ion at *m/z* 343.1396 yielded the fragment at *m/z* 299.1506 ($C_{15}H_{23}O_6^-$) corresponding to a neutral loss of *m/z* 44 that implied the loss of CO_2 suggesting the presence of a carboxylic acid group. The presence of a carbonyl at C-7 was further confirmed from the ^{1}H NMR chemical shifts of the nearest protons, H-2 and H-6 that appear to be more deshielded in **34** compared to **31**, that is, 6.77 and 2.31/2.20 ppm compared to 5.85 and 2.13/2.13 ppm for H-2 and H-6a/H-6b in **34** and **31**, respectively. Thus, compound **34** was newly characterised as

Table 1
^{1}H and ^{13}C NMR spectroscopic data of compound **34** (δ in ppm, J in Hz).

No.	**31**		**34**	
	δ_H	δ_C	δ_H	δ_C
1		136.6		*
2	5.85, m	130.0	6.77, m	141.9
3a	2.24, m	33.2	2.38, m	33.7
3b	2.03, m		2.15, m	
4	2.24, m	39.0	2.31, m	38.0
5a	1.88, m	30.1	1.93, m	29.6
5b	1.51, m		1.50, m	
6a	2.13, m	28.4	2.31, m	27.2
6b	2.13, m		2.20, m	
7a	4.21, d (12.3)	76.5		*
7b	4.14, d (12.3)			
8		155.9		151.6
9a	5.06, m	111.2	5.19, s	115.3
9b	4.97, m		5.06, s	
10a	4.12, s	66.7	4.43, d (12.7)	74.3
10b	4.12, s		4.24, d (12.7)	
1′	4.39, d (8.1)	103.3	4.45, d (8.3)	104.2
2′	3.26, dd (9.2, 8.1)	76.2	3.28, dd (9.2, 8.3)	76.3
3′	3.43, m	79.1	3.46, dd (9.2, 8.8)	79.1
4′	3.38 – 3.35, m	72.7	3.37, m	72.7
5′	3.38 – 3.35, m	78.9	3.41, m	79.0
6′a	3.89, dd (12.3, 1.7)	63.8	3.90, dd (12.3, 2.0)	63.9
6′b	3.70, dd (12.3, 5.5)		3.71, dd (12.3, 5.5)	

Data collected at 600 MHz for ^{1}H and 150 MHz for ^{13}C NMR in D_2O:CD_3OD (4:1). Spectra were referenced to TSP-d_4 (0.01 % w/v) at δ_H 0.00. *Could not be seen in the HMBC spectrum.

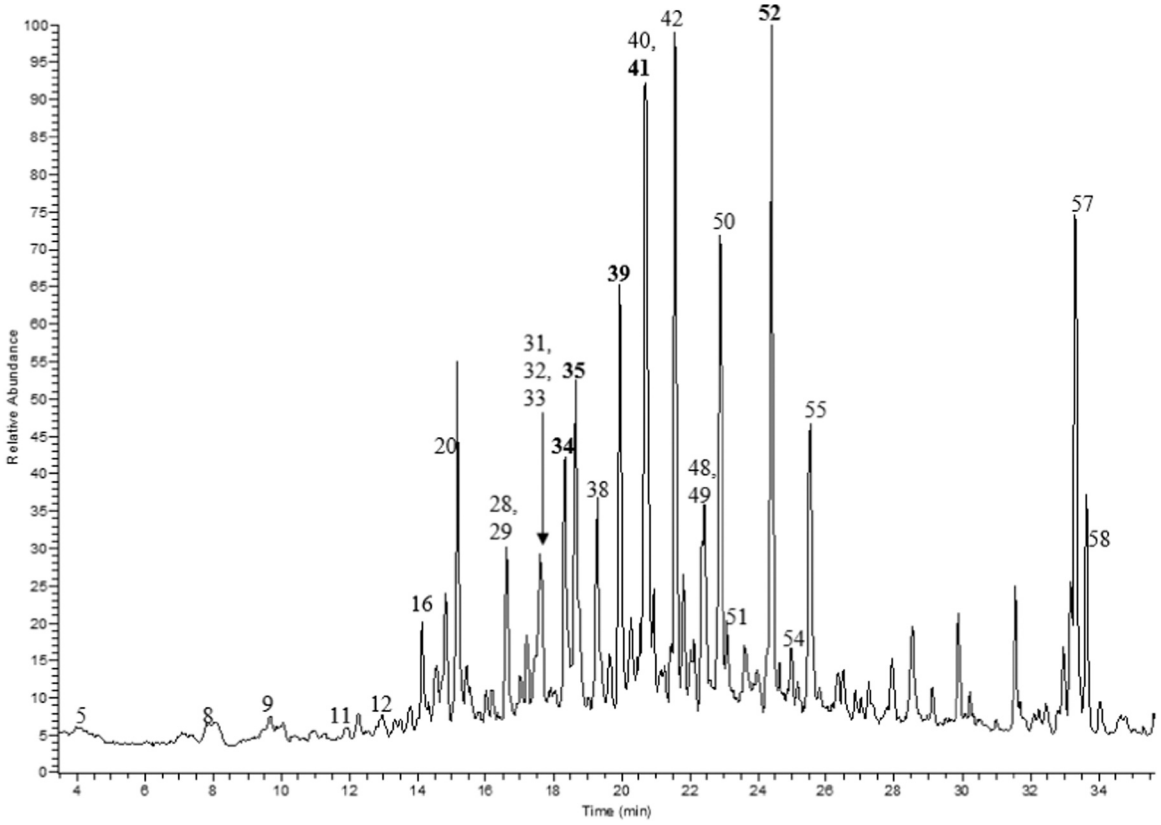

Fig. 2. Total ion chromatogram (negative ion mode) of *A. leucotrichus* seeds extracted with water/methanol (4:1). Labelled peaks represent some of the isolated and identified compounds. Numbers in bold correspond to previously undescribed compounds. Peak numbers relate to structure numbers in Fig. 1 and to the full data in Table S1.

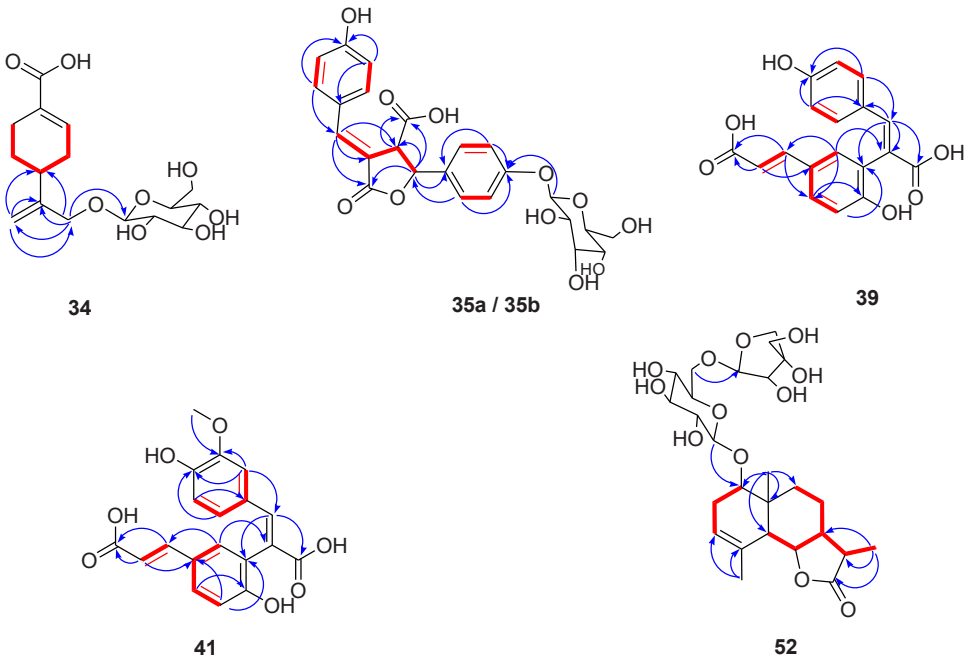

Fig. 3. Key selected HMBC (arrows) and ^{1}H-^{1}H COSY (red bonds) correlations of compounds **34**, **35**, **39**, **41** and **52**.

10-hydroxylimonene-7-oic acid-10-*O*-*β*-glucopyranoside. Too little was isolated for further investigative work towards the absolute configuration of **34**, but given that *R*-perillaldehyde has been shown to be a major metabolite in ethanol extracts of *A. leucotrichus* (Bellau et al., 2022), we can assume that the *R* chirality at C1 has been maintained in the new derivative **34**, and also the related **31** which both appear to be biosynthetically closely related to perillaldehyde.

Compound **35** was a mixture of isomers (**35a** and **35b)** that partially separated across two sequential HPLC fractions. However, UHPLC-MS of both fractions showed the same deprotonated ion peak at exactly the same retention time of 18.65 min. Based on this ion at *m/z* 487.1243 [M-H]$^{-}$ (calcd. for $C_{24}H_{23}O_{11}^{-}$, 487.1246), the compound was given the molecular formula of $C_{24}H_{24}O_{11}$. However, the major ion peak in the MS spectrum was *m/z* 443.1343 ($C_{23}H_{23}O_{9}^{-}$) which corresponded to the fragment formed after a neutral loss of *m/z* 44 indicating the presence of a terminal carboxylic acid group in both molecules. The MS/MS of both compounds showed the major fragment at *m/z* 237.0925 ($C_{16}H_{13}O_{2}^{-}$) formed after further neutral losses of *m/z* 162 and *m/z* 44 from the ion at *m/z* 443.1343 ($C_{23}H_{23}O_{9}^{-}$) indicating the presence of a hexose and a second carboxylic acid group in the compounds. From the ^{1}H NMR spectra of both fractions, it was observed that each contained two sets of signals in unequal proportions. Based on the proportion of signals, the earlier smaller fraction mainly contained **35b** with traces of **35a** whilst the reverse was observed for the second bigger fraction indicating that the two compounds could not be successfully separated on the preparative HPLC. Unfortunately, the earlier fraction was not strong enough for 2D-NMR whereas the second fraction showed good NMR data for both compounds in the mixture. Based on the signals in the ^{1}H NMR spectrum of this sample, the two compounds were in a ratio of 5:1 (**35a:35b**). Structural elucidation of both compounds was therefore done using the NMR data of this second fraction.

Considering the set of major signals corresponding to compound **35a** in the ^{1}H NMR spectrum, the downfield region showed two pairs of signals at δ_H 7.56 (2H, d, J = 8.8 Hz, H-2′/6′), 6.95 (2H, d, J = 8.8 Hz, H-3′/5′) and 7.35 (2H, d, J = 8.8 Hz, H-2/6), 7.15 (2H, d, J = 8.8 Hz, H-3/5) with the signals in each pair correlating in the ^{1}H-^{1}H COSY spectrum and belonging to the same spin system in the ^{1}H-^{1}H TOCSY spectrum. This confirmed the presence of two *para*-disubstituted aromatic rings. Both phenyl groups were part of a phenylpropanoid moiety as judged by HMBC interactions from H-2′/H-6′ to C-7′ (δ_C 143.5) and from the olefinic proton at δ_H 7.66 (1H, d, J = 2.0 Hz, H-7′) to a carbonyl at δ_C 178.1 (C-9′). Likewise, the HMBC spectrum evidenced cross peaks from H-2/H-6 to an aliphatic-like resonance at δ_C 86.1 (C-7) and from H-7 to C-8 (δ_C 59.4) and C-9 (δ_C 179.9). The two aliphatic-like protons at δ_H 5.76 (1H, d, J = 3.0 Hz, H-7) and 4.02 (1H, dd, J = 3.0, 2.0 Hz, H-8) in the ^{1}H NMR spectrum featured additional HMBC cross peaks from H-7 to C-9′ and from H-8 to C-7′ and C-8′ (δ_C 123.9). These couplings of H-7 and H-8 combined with an additional HMBC correlation from H-7′ to C-8 suggested the presence of a *γ*-lactone ring in **35a**. These data are characteristic of an 8–8′ lignan-type of compound. As in compound **34**, the patterns in the coupling constants of the sugar protons supported a *β*-glucopyranose with an anomeric signal at δ_H 5.09 (1H, d, J = 7.6 Hz, H-1″) in the ^{1}H NMR spectrum. The sugar was deduced to be attached at C-4 on the lignan backbone according to HMBC cross peaks from H-1″ and H-2/H-6 to C-4 (δ_C 160.1). The double bond was suggested to be in an *E*-configuration based on the small coupling constant (J = 2.0 Hz) of H-7′ from its long-range coupling to H-8 (Dall'Acqua et al., 2004; Miao et al., 2022). Comparing this data to that reported in literature shows that the aglycone part of **35a** is similar to the lignan derivative initially isolated from the roots of *Chaerophyllum hirsutum* (Dall'Acqua et al., 2004), and more recently from *Ligusticum chuanxiong* Hort (Miao et al., 2022) and also identified in this study (peak **51**). The small coupling constant of 3.0 Hz between H-7 and H-8 as well as a strong NOE correlation between themselves, and a NOE correlation of both signals to H-2/6 (Fig. 4) seems to suggest that the two hydrogens are in a *cis*-configuration to each other which would suggest a 7 *R*, 8 *R* or 7*S*, 8*S* configuration (Dall'Acqua et al., 2004).

The set of minor signals corresponding to compound **35b** had the same pattern as those of **35a** with very similar chemical shifts except for H-7′ and H-8 which appeared to be slightly more downfield in **35a**, that is, 7.66 and 4.02 ppm for H-7′ and H-8 respectively in **35a** compared to 7.24 and 3.90 ppm in **35b** (Table 2). The signals of both compounds also showed similar correlations in all the 2D spectra, including the same NOE correlations, suggesting similar structures and implying that compound **35b** was a close structural isomer of **35a**. The data suggests that the two compounds are diastereomers, that are marginally separated under a single HPLC peak. Despite the shift difference of H-8 in the two isomers, the presence of NOE correlations between H-7 and H-8 in

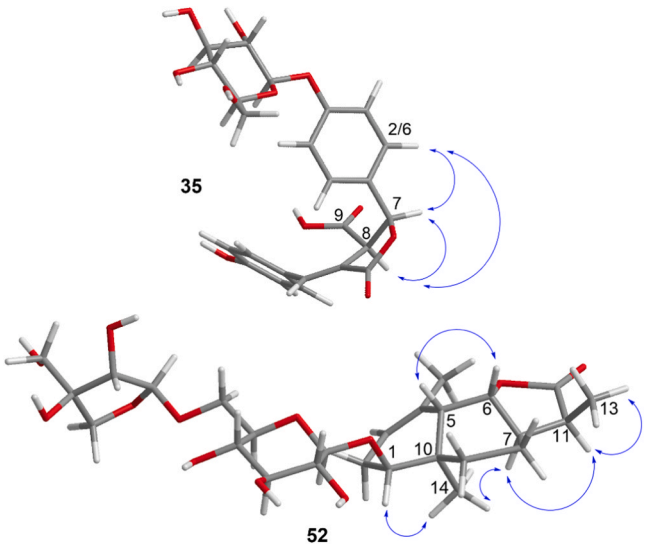

Fig. 4. Key NOESY correlations of compounds **35** and **52**.

Table 2
^{1}H and ^{13}C NMR spectroscopic data of compounds **35a** and **35b** (δ in ppm, J in Hz).

No.	**35a**		**35b**	
	δ_H	δ_C	δ_H	δ_C
1		137.4		136.7
2	7.35, d (8.8)	130.2	7.41, d (8.7)	130.5
3	7.15, d (8.8)	120.0	7.18, d (8.7)	119.8
4		160.1		160.1
5	7.15, d (8.8)	120.0	7.18, d (8.7)	119.8
6	7.35, d (8.8)	130.2	7.41, d (8.7)	130.5
7	5.76, d (3.0)	86.1	5.79, d (5.3)	85.2
8	4.02, dd (3.0, 2.0)	59.4	3.90, m	59.8
9		179.9		179.9
1′		129.1		129.1
2′	7.56, d (8.8)	136.0	7.78, d (8.7)	136.0
3′	6.95, d (8.8)	118.9	6.92, d (8.7)	118.6
4′		161.5		160.8
5′	6.95, d (8.8)	118.9	6.92, d (8.7)	118.6
6′	7.56, d (8.8)	136.0	7.78, d (8.7)	136.0
7′	7.66, d (2.0)	143.5	7.24, d (2.0)	145.4
8′		123.9		128.7
9′		178.1		173.6
1″	5.09, d (7.6)	103.2	5.11, d (7.6)	103.2
2″	3.53, dd (9.3, 7.6)	76.0	3.55, dd (9.2, 7.6)	76.2
3″	3.56, m	78.9	3.58, m	78.9
4″	3.47, dd (9.3, 9.0)	72.5	3.49, dd (9.2, 9.0)	72.7
5″	3.59, ddd (9.0, 5.6, 2.3)	78.9	3.62, m	78.9
6″a	3.90, dd (12.5, 2.3)	63.6	3.91, dd (12.5, 2.1)	63.6
6″b	3.73, dd (12.5, 5.6)		3.73, dd (12.5, 5.5)	

Data collected at 600 MHz for ^{1}H and 150 MHz for ^{13}C NMR in D_2O:CD_3OD (4:1). Spectra were referenced to TSP-d_4 (0.01 % w/v) at δ_H 0.00.

both suggests that if there is a change in orientation, it occurs at both positions. A flip in the orientation of both H-7 and H-8 also changes the orientation of the carbonyl group at C-9 which could in turn affect the chemical shift of H-7′. The data suggests that one of the compounds is 7 *R*, 8 *R* and the other 7*S*, 8*S*. Besides, the two configurations are feasible from the biogenesis point of view as the oxidative coupling involved in lignan formation may involve both the upward and downward orientation of the intermediate radical species.

As the diastereomeric mixture of **35a** and **35b** could not be obtained in a sufficient amount to confirm their absolute configurations, we are unable to define the exact stereochemistry at C7/8 in each, or indeed to test whether the stereochemistry at C8 is interconvertible by enolisation. We suggest that **35** be named ammodauside. Related lignans with methoxylation in the 3-position known as ligusticumacid E (7 *R*, 8 *R*) or chuanxioside C (7*S*, 8*S*) have been previously isolated from the rhizomes of *Ligusticum chuanxiong* (Yuan et al., 2020; Zhang et al., 2018). Indeed, one of these methoxylated derivatives was also identified in this study (peak **38**) (Table S1 and List S1 in supplementary material).

Compound **39** was assigned a molecular formula of $C_{18}H_{14}O_6$ based on its deprotonated ion at *m/z* 325.0718 [M-H]$^-$ (calcd. for $C_{18}H_{13}O_6^-$, 325.0718). Its MS/MS data included a fragment at *m/z* 281.0823 ($C_{17}H_{11}O_4^-$) resulting from the loss of CO_2 from a terminal carboxylic acid group. The major fragment was at *m/z* 237.0924 ($C_{16}H_{13}O_2^-$) from the loss of a second CO_2 molecule indicating that the compound contained two carboxylic acid groups. The ^{1}H NMR data (Table 3) included signals at δ_H 7.03 (2H, d, J = 8.6 Hz, H-2/6) and 6.69 (2 H, d, J = 8.6 Hz, H-3/5) showing the presence of a *para*-disubstituted benzene ring of the AA′BB′ spin type. The presence of an additional ABX spin system was shown by the aromatic signals at δ_H 7.55 (1H, dd, J = 8.5, 2.2 Hz, H-6′), 7.27 (1H, d, J = 2.2 Hz, H-2′) and 6.99 (1H, d, J = 8.5 Hz, H-5′). The two signals at δ_H 7.37 (1H, d, J = 15.9 Hz, H-7′) and δ_H 6.29 (1H, d, J = 15.9 Hz, H-8′) indicated the presence of a double bond in the *E*-configuration. An additional olefinic hydrogen was observed at δ_H 7.67 (1H, s, H-7). The presence of two carboxylic acid groups was confirmed by the HMBC cross peaks from H-7′ and H-8′ to a carbonyl at δ_C 178.4 (C-9′) and from H-7 to a second carbonyl at δ_C 178.6 (C-9). The HMBC coupling patterns across the molecule (Fig. 3) were consistent with the presence of two phenylpropanoid units that were bonded based on HMBC cross peaks from H-7 to C-3′ (δ_C 129.2) and from H-2′ to C-8 (δ_C 134.8) which is characteristic of an 8–3′ neolignan. The ^{13}C chemical shifts of the other quaternary carbons were also obtained using their HMBC correlations. These included couplings of C-1 with H-7 and H-3/5, C-4 with H-3/5 and H-2/6, C-8 with H-7 and H-2′, C-1′ with H-5′ and H-8′, C-3′ with H-7 and H-5′, and C-4′ with H-2′, H-5′, and H-6′. The C7-C8 double bond was determined to be in an *E*-configuration based on a NOESY correlation that was observed between H-2′ and H-2/6 which is only possible if the bond is in an *E*-configuration. Additionally, the chemical shift of H-7 is the same as that of H-7′ in compound **35a** (7.66 ppm) and compound **38** (7.67 ppm) both in which the double bond is in an *E*-configuration. Thus, the structure of **39** was newly characterised and trivially named ammodaucic acid A. This compound has been previously synthesised and the reported NMR data is in agreement with our study (Qiu-An et al., 2008). However, this is the first time it has been reported from plants or any natural source.

Compound **41** was assigned a molecular formula of $C_{19}H_{16}O_7$ based

Table 3
^{1}H and ^{13}C NMR spectroscopic data of compounds **39** and **41** (δ in ppm, J in Hz).

No.	**39**		**41**	
	δ_H	δ_C	δ_H	δ_C
1		130.7		130.7
2	7.03, d (8.6)	134.7	6.56, d (1.8)	115.3
3	6.69, d (8.6)	118.2		150.1
4		159.1		149.5
5	6.69, d (8.6)	118.2	6.80, d (8.2)	118.3
6	7.03, d (8.6)	134.7	6.88, dd (8.2, 1.8)	128.7
7	7.67, s	140.7	7.71, s	143.0
8		134.8		139.8
9		178.6		177.0
1′		131.0		130.9
2′	7.27, d (2.2)	134.0	7.34, d (2.2)	134.3
3′		129.2		131.0
4′		158.1		159.1
5′	6.99, d (8.5)	119.8	7.03, d (8.5)	119.6
6′	7.55, dd (8.5, 2.2)	131.8	7.57, dd (8.5, 2.2)	132.7
7′	7.37, d (15.9)	144.1	7.46, d (15.9)	146.1
8′	6.29, d (15.9)	123.9	6.30, d (15.9)	121.6
9′		178.4		176.8
OMe			3.39, s	58.1

Data collected at 600 MHz for ^{1}H and 150 MHz for ^{13}C NMR in D_2O:CD_3OD (4:1). Spectra were referenced to TSP-d_4 (0.01 % w/v) at δ_H 0.00.

on its deprotonated ion at *m/z* 355.0822 [M-H]$^-$ (calcd. for $C_{19}H_{15}O_7^-$, 355.0823). The molecular mass of **41** differs from that of compound **39** by 30 Da suggesting that **41** probably differs from **39** by having an additional methoxy group. As in **39**, the presence of two carboxylic groups in **41** was revealed by the MS/MS data that showed fragments corresponding to two successive neutral losses of CO_2 at *m/z* 311.0928 ($C_{18}H_{15}O_5^-$) and 267.1031 ($C_{17}H_{15}O_3^-$), with the fragment from the second loss also representing the major fragment. An additional fragment at *m/z* 252.0796 ($C_{16}H_{12}O_3^-$) corresponded to a loss of the methyl group from the product ion at *m/z* 267.1031 ($C_{17}H_{15}O_3^-$) in line with the suggested presence of a methoxy group. The ^{1}H NMR spectrum of **41** showed the presence of an ABX spin system, a double bond in an *E*-configuration, an olefinic proton and a methoxy signal at δ_H 3.39 (3H, s). However instead of an AA′BB′ system as in **39**, the data confirmed the presence of a second ABX system as revealed by the signals at δ_H 6.56 (1H, d, J = 1.8 Hz, H-2), 6.80 (1H, d, J = 8.2 Hz, H-5) and 6.88 (1H, dd, J = 8.2, 1.8 Hz, H-6). The HMBC spectrum showed a coupling of the methoxy protons to C-3 indicating that they are attached to this position. Besides the couplings to C-3, the other correlations observed in the HMBC spectrum were similar to those of **39** (Fig. 3) consistent with the identification of **41** as an 8–3′ neolignan with the same base structure as **39**, but with an additional methoxy group. Instead of two *p*-coumaric acid units as in **39**, **41** consists of a ferulic acid, and a *p*-coumaric unit joined at the 8- and 3′ positions of the former and latter respectively. Thus, the structure of **41** was newly characterised and trivially named ammodaucic acid B. This is the first report of this compound from any source. However, a ferulic acid dimer with a similar structure has been isolated from extracts of saponified cell walls of cocksfoot, switchgrass and suspension-cultured corn, along with other dimers with different coupling patterns (Ralph et al., 1994). This same ferulic acid dimer was a substructure of a ferulic acid trimer isolated from maize bran (Bunzel et al., 2005).

Compound **46** was assigned a molecular formula of $C_{26}H_{30}O_{12}$ based on its deprotonated ion at *m/z* 533.1661 [M-H]$^-$ (calcd. for $C_{26}H_{29}O_{12}^-$, 533.1665). The major fragment in its MS/MS was at *m/z* 353.1035 ($C_{20}H_{17}O_6^-$) formed after the loss of $C_6H_{12}O_6$ which indicated the presence of a hexose in the compound. Additional fragments at *m/z* 338.0802 ($C_{19}H_{14}O_6^-$) and 323.0563 ($C_{18}H_{11}O_6^-$) indicated successive losses of two methyl groups from the major fragment. Another fragment at *m/z* 325.1085 ($C_{19}H_{17}O_5^-$) corresponded to the loss of a -CO group. Attempts to obtain a pure fraction of the compound for complete NMR analysis were unsuccessful as the ^{1}H NMR spectrum of the fraction containing the compound was dominated by signals of another compound that elutes closely on the HPLC. It was therefore not possible to get exploitable 2D spectra for compound **46** for complete elucidation. However, the ^{1}H NMR signals for the compound were clear enough to facilitate a putative identification of its structure. These included δ_H 9.47 (1H, d, J = 8.0 Hz) which suggested the presence of an aldehyde group which was conjugated to an *E*-configured double bond as indicated by the signals at δ_H 7.62 (1H, d, J = 15.7 Hz) and 6.68 (1H, dd, J = 15.7, 8.0 Hz). The aromatic region indicated the presence of an ABX aromatic system from the signals at δ_H 7.66 (1H, d, J = 1.9 Hz), 7.66 (1H, dd, J = 8.8, 1.9 Hz) and 6.91 (1H, d, J = 8.8 Hz) as well as a tetra-substituted aromatic ring with two *meta*-coupled protons shown by the signals at δ_H 7.27 (1H, d, J = 1.8 Hz) and 7.06 (1H, d, J = 1.8 Hz). Additionally, the two signals at δ_H 5.46 (1H, dd, J = 8.4, 5.0 Hz) and 4.65 (1H, dd, J = 10.6, 8.4 Hz) suggested a missing unobserved signal with couplings of *ca.* 10.6 and 5.0 Hz that was potentially overlapping with other signals. These three signals would indicate the same C7-C8 connection as observed for compounds **39** and **41**, but with the bond being saturated in **46**. An anomeric signal at δ_H 4.45 (1H, d, J = 7.9 Hz) indicated the presence of a *β*-hexose whereas the two signals at δ_H 3.91 (3H, s) and 3.87 (3H, s) suggested the presence of two methoxy groups in the compound. Combining the UHPLC-MS and ^{1}H NMR data suggests that compound **46** could be a dimethoxylated glycoside derivative of compound **39** in which the carboxylic acid group in 9′-position has been reduced to an aldehyde group. With the lack of carbon data, this structure could not be confirmed, but **46** would otherwise also be a previously undescribed compound.

Compound **52** was assigned a molecular formula of $C_{26}H_{40}O_{12}$ based on its deprotonated ion at *m/z* 543.2436 [M-H]$^-$ (calcd. for $C_{26}H_{39}O_{12}^-$, 543.2447). Its MS/MS data showed a fragment at *m/z* 499.2546 ($C_{25}H_{39}O_{10}^-$) from the loss of CO_2 from the [M-H]$^-$ ion. Additional fragments at *m/z* 411.2037 ($C_{21}H_{31}O_8^-$) and *m/z* 231.1394 ($C_{15}H_{19}O_2^-$) indicated successive losses of a pentose fragment and a hexose from the [M-H]$^-$ ion. The ^{1}H NMR spectrum (Table 4) showed no aromatic signals but (excluding the sugar signals) evidenced the presence of three methyls at δ_H 0.91 (3H, s, H-14), 1.13 (3H, d, J = 7.5 Hz, H-13) and 1.82 (3H, m, H-15); three methylenes at δ_H 2.16 (1H, br.d, J = 19.0 Hz, H-2a)/2.30 (1H, dt, J = 19.0, 2.6 Hz, H-2b), 1.32 (1H, td, J = 13.0, 13.0, 6.5 Hz, H-8a)/1.64 (1H, tt, J = 13.0, 6.5 Hz, H-8b), and 1.22 (1H, ddd, J = 13.0, 11.0, 6.5 Hz, H-9a)/2.25 (1H, ddd, J = 13.0, 6.5, 1.2 Hz, H-9b); and six methines at δ_H 3.56 (1H, dd, J = 3.0, 1.4 Hz, H-1), 5.45 (1H, br.s, H-3), 2.39 (1H, d, J = 9.7 Hz, H-5), 4.83 (1H, dd, J = 9.7, 6.5 Hz, H-6), 2.89 (1H, m, H-7), and 3.07 (1H, dq, J = 7.5, 7.5 Hz, H-11) with H-3 being an olefinic proton. Its HMBC spectrum exhibited cross peaks from H-13 to C-11 (δ_C 41.2), C-7 (δ_C 40.5) and a carbonyl at δ_C 187.5 (C-12); from H-14 to C-9 (δ_C 32.5), C-10 (δ_C 37.6), C-1 (δ_C 83.3) and C-5 (δ_C 44.0); and from H-15 to C-4 (δ_C 137.4) and C-3 (δ_C 123.0). Additionally, the ^{1}H-^{1}H-COSY and ^{1}H-^{1}H TOCSY spectra evidenced two main spin systems in **52**, mainly between H-5, H-6, H-7, H-8a, H-8b, H-9a, H-9b, H-11 and H-13; and between H-1, H-2a, H-2b, H-3 and H-15 consistent with the presence of a eudesmane-type sesquiterpene lactone (Han et al., 2005; Kurimoto et al., 2012). The ^{1}H NMR spectrum also showed the presence of an anomeric signal at δ_H 4.46 (1H, d, J = 8.0 Hz, H-1′) which together with its associated signals showed the presence of a *β*-glucopyranose. The HMBC cross peak from H-1′ to C-1 showed that the

Table 4
^{1}H and ^{13}C NMR spectroscopic data of compounds **52** (δ in ppm, J in Hz).

No.	δ_H	δ_C
1	3.56, dd (3.0, 1.4)	83.3
2a	2.16, br,d (19.0)	30.5
2b	2.30, dt (19.0, 2.6)	
3	5.45, br,s	123.0
4		137.4
5	2.39, d (9.7)	44.0
6	4.83, dd (9.7, 6.5)*	85.2
7	2.89, m	40.5
8a	1.32, td (13.0, 13.0, 6.5)	27.5
8b	1.64, tt (13.0, 6.5)	
9a	1.22, ddd (13.0, 11.0, 6.5)	32.5
9b	2.25, ddd (13.0, 6.5, 1.2)	
10		37.6
11	3.07, dq (7.5, 7.5)	41.2
12		187.5
13	1.13, d (7.5)	12.3
14	0.91 s	25.2
15	1.82, m	24.2
1′	4.46, d (8.0)	103.9
2′	3.23, dd (9.0, 8.0)	76.2
3′	3.44, m	78.8
4′	3.40, m	72.7
5′	3.49, ddd (9.6, 5.8, 2.0)	78.2
6′a	3.98, dd (11.6, 2.0)	70.3
6′b	3.72, dd (11.6, 5.8)	
1″	5.12, d (2.9)	111.9
2″	3.97, d (2.9)	79.6
3″		82.4
4″a	4.01, d (10.1)	76.6
4″b	3.87, d (10.1)	
5″a	3.65, d (3.7)	66.7
5″b	3.65, d (3.7)	

Data collected at 600 MHz for ^{1}H and 150 MHz for ^{13}C NMR in D_2O:CD_3OD (4:1). Spectra were referenced to TSP-d_4 (0.01 % w/v) at δ_H 0.00. *overlapped with residual water signal.

β-glucopyranose was attached to the hydroxyl at C-1 of the sesquiterpene lactone aglycone.

Comparing the data of the sesquiterpene lactone glycoside identified so far (only part of the compound) to the literature shows that it is similar to that of eudesma-3-enolide 1-*O*-β-glucopyranoside, a compound that has been previously isolated from *Sonchus transcaspicus* (Han et al., 2005). However, the ^{1}H NMR spectrum additionally showed a second anomeric signal at δ_H 5.12 (1H, d, J = 2.9 Hz, H-1″) which together with its associated signals, a δ_C value of 111.9 ppm for C-1″ and the appearance of C-3″ as a quaternary carbon at δ_C 82.4 indicated the presence of a β-apiofuranose. The appearance of the C-6′ of the glucose at a significantly downfield chemical shift of 70.3 ppm together with a coupling observed between H-6′b of the glucose and C-1″ of the apiose in the HMBC confirmed that the apiose was attached to the 6-position of the glucose. The relative configurations of **52** around the decalin ring were determined by examination of J couplings alongside COSY and NOESY data. Accordingly, the orientations of H-1 and the methyl H-14 were supported by strong NOE interactions from H-7 to H-14, and then from H-14 to H-1. In addition, the lack of NOE interaction between H-14 and H-5 indicated a *trans*-decalin in **52**. The methyl H_3-13 did not show any NOE interaction except that expected with H-11. Conversely, H-11 also had an NOE with H-7α, indicating a *cis* relationship of these hydrogens. Together, these interactions indicate that the 13-methyl group is β-orientated as shown. Compound **52** is novel and trivially named ammodolide. Although free apiose is rare in plants, apiofuranose is commonly seen as a module in glycosylated natural products and it has been suggested that since the discovery of apiin (a well-known apiosylated flavonoid glycoside), over 1000 apiosylated conjugates have been recorded in plants (Picmanova and Moller, 2016).

Four indole alkaloid derivatives including non-glycosylated (**8**) and glycosylated (**5**, **7**, and **9**) compounds were also isolated. However, their NMR data was not sufficient for complete elucidation of their structures. Nevertheless, this seems to be the first report of indole alkaloid derivatives in *A. leucotrichus*. Of the known compounds isolated in this study, some have been identified from *A. leucotrichus* in previous studies. Ammolactone-A (**58**) and 3-hydroxyperillaldehyde (see supplementary material for spectroscopic data) were previously identified from the seeds bought from local markets in N. Africa (Muckensturm et al., 1997; Bellau et al., 2022). The stereostructure for ammolactone-A (**58**) has been revised (Hernandez and Joseph-Nathan, 2000) and this revision is depicted in Fig. 1. Similarly, vanillin (**26**) and *p*-coumaric acid (**30**) were identified from the aerial parts of the plant collected from Algeria (Mouderas et al., 2020). Apigenin 6,8-di-C-β-glucopyranoside (**33**), luteolin 7-O-β-glucopyranoside (**40**) and malonyl dicaffeoyl quinic acid (**47**) have also been previously detected in the aerial parts (Ziani et al., 2019). The same study also putatively identified two isomers of luteolin-*O*-(malonyl-hexoside), one of which could correspond to luteolin 7-*O*-(6″-*O*-malonyl)-β-glucopyranoside (**50**) identified in our study. Apigenin 7-*O*-β-glucopyranoside (**48**), balanophonin 4-*O*-β-glucopyranoside (**49**), chrysoeriol (**54**), luteolin (**55**), ammodaucine A (**57**), and ammodaucine D (**59**) were identified from fruits collected from Morocco (Hajib et al., 2023). Isoleucine (**3**), guanosine (**4**), tryptophan (**10**), salidroside (**13**), olivil *O*-β-glucopyranoside (**23** or **32**), licoagroside B (**29**) and secoisolariciresinol dihexoside (**37**) have also been putatively identified from fruits harvested from Algeria (Abderrezag et al., 2021). Additionally, scopolin (**22**) was identified from fruits of spontaneous plants growing in Algeria (Abderrezag et al., 2024).

Terpene and lignan derivatives constituted the previously undescribed compounds as well as several of the identified known compounds. Terpene derivatives have commonly been identified in the essential oil of *A. leucotrichus* especially *R*-perillaldehyde and limonene that have frequently been identified as the major compounds in the plant's oil (Idm'hand et al., 2020). *R*-Perillaldehyde and ammolactone-A were major metabolites in ethanol extracts (Bellau et al., 2022). Very few lignan derivatives have also been identified from the plant (Hajib et al., 2023).

Terpenes have applications in food, medicine and cosmetic industries with reports of potential pharmacological effects including anticancer, and anti-inflammatory activities (Fan et al., 2023; Masyita et al., 2022). They also have bioprotectant activity against insect pests, nematodes and other phytopathogens (Isman, 2000). The aglycone of **34** is similar to perillic acid, a substance with potential antidiabetic activity that has been described as an effective preservative and antimicrobial agent (Keserla et al., 2023; Rehdorf and Kleber, 2017). Previous works on eudesmane-type sesquiterpenes similar to **52** indicate they have the potential of being used in the formulation of anti-cancer (Wang et al., 2023) and anti-inflammatory drugs (Dang et al., 2019), as well as herbicides (Wu et al., 2022). Lignans have known antioxidant properties and have been shown to contribute to important biological effects including antiviral and antibacterial effects, as well as playing an ecological function by helping plants to counteract the effects of both biotic and abiotic stresses (Šamec et al., 2021). Phellinsin A, a lactone-type lignan similar to **35**, **38** and **51**, has shown great potential to prevent and treat cardiovascular diseases such as hyperlipemia and arteriosclerosis (Kim et al., 2006). This study has aimed to characterise the polar metabolome of *A. leucotrichus*, via a metabolomics approach, designed to rapidly annotate known metabolites but also to provide data on new metabolites, in order to facilitate future investigations to determine the active principles behind the recorded traditional medicinal uses of this plant. In this context, the biological effects of terpenes and lignans highlighted may be important in the reported traditional use and potentially point to other possible uses of *A. leucotrichus*. However, to explore this, a larger scale chromatographic isolation or bioassay-guided fractionation would be necessary to obtain sufficient quantities of each compound for further biological assessment. As *A. leucotrichus* is available from commercial sources, it may also be advantageous to rapidly compare multiple samples, via the profiling technology described, in order to relate profiles to biological efficacy and thus to better target individual compounds for more detailed biomedical work.

3. Conclusions

In summary, the profile of identified compounds mainly comprised phenolic acid derivatives and lignans. These mainly included derivatives of ferulic and *p*-coumaric acids, for both the known and newly characterised (**35**, **39**, **41** and **46**) compounds. Relatedly, most of the known lignans that were characterised are based on coniferoyl alcohol as the base monolignol. The other compounds included flavonoid, terpene, phenolic and alkaloid derivatives together with some amino- and organic acids. Interestingly, despite being very common secondary metabolites, only eight flavone derivatives were isolated and identified from the extract. Besides the lignan derivatives, the other two new compounds (**34** and **52**) described are terpene derivatives.

4. Experimental

4.1. General experimental procedures

Acquisition of UHPLC-MS data was done using an LTQ-Orbitrap Elite mass spectrometer connected to a Dionex UltiMate 3000 RS UHPLC system (Fisher Scientific). A reversed-phase Hypersil GOLD C18 selectivity HPLC column (3 μm, 30 ×2.1 mm i.d. Thermo Fisher Scientific) kept at 35 °C was loaded with samples (10 μL). The solvent system was composed of water + 0.1 % formic acid (A) and acetonitrile + 0.1 % formic acid (B). The elution gradient was as follows: 0–5 min, 0 % B; 5–27 min, 31.6 % B; 27–34 min, 45 % B; and 34–37.5 min, 75 % B. The total run time was 40 min using a flow rate of 0.3 mL/min. A heated ESI device was used to collect mass spectra, which were acquired in negative mode with a resolution of 120,000 spanning the m/z 50–1500 range. The capillary temperature, source voltage, sheath gas, auxiliary gas, and

sweep gas were each set to 350°C, 2.5 kV, 35 (arbitrary units), 10, and 0.0 (arbitrary units), respectively. Other acquisition parameters were set to their default values. Using an isolation width of m/z 2, automatic MS/MS fragmentation of the top four ions was carried out. High-energy C-trap dissociation (HCD) was used to fragment ions, with a normalised collision energy of 65 and an activation time of 0.1 ms. Xcalibur v. 2.2 (Thermo Fisher Scientific) was used to collect and examine the data.

For NMR data, a Bruker Avance 600 MHz NMR spectrometer (Bruker Biospin, Germany) was used to collect ^{1}H and ^{13}C NMR spectra at 600.05 MHz (^{1}H) and 150 MHz (^{13}C) respectively. One dimensional ^{1}H spectra were acquired with a 5 mm TCI cryoprobe using a zgpr pulse sequence with a 90°angle. A pre-saturation delay of 5 s was used to suppress the residual water signal. Spectra had a spectral width of 12 ppm and 64,000 data points. Transformation of the FIDs was automatically done using Topspin version 4.2.0 using an exponential window and a line broadening of 0.5 Hz. ^{1}H–^{1}H correlation spectroscopy (COSY) spectra were acquired using the cosyprqf pulse sequence at a frequency of 600.05 MHz in both dimensions, with acquisition times of 0.1434 and 0.0896 s. The sweep widths were 7142.9 Hz, and 32 transients were used to acquire 1024 data points in each dimension. ^{1}H–^{13}C heteronuclear single quantum coherence (HSQC) spectra were acquired using the hsqcetgpsi2 pulse sequence, at 600.05 and 150.9 MHz, with acquisition times of 0.1433 and 0.00212 s. Sweep widths of 7142.9 Hz and 30,120.5 Hz were used to obtain the data. Using 128 transients, 2048 and 1024 data points were acquired. The ^{1}H–^{13}C heteronuclear multiple bond correlation (HMBC) spectra were acquired using the hmbcgpndqf pulse sequence with the same acquisition parameters as for the HSQC data. All the spectra were obtained at 300 K in D_2O:CD_3OD (4:1), and chemical shifts are given in δ, relative to 3-(trimethylsilyl) propionic acid-d_4, 0.01 % *w/v* (TSP-d_4) standard. Phasing and baseline correction were carried out within the instrument software and data was analysed using MestreNova software.

4.2. Plant material

Dried seeds of *A. leucotrichus* harvested from Marrakech in western Morocco were sourced from Arij Al Ghaba, a local certified co-operative. Species identification was carried out by representatives from the co-operative. The samples were 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 use.

4.3. Extraction and fractionation

For preliminary screening of the crude extract, three replicates (15 mg) of the milled sample were suspended in either H_2O/CH_3OH (4:1 v/v, 1 mL) for UHPLC-MS analysis or in an NMR solvent consisting of D_2O/CD_3OD (4:1 v/v) containing 3-(trimethylsilyl) propionic acid-d_4, 0.01 % *w/v* (TSP-d_4) (1 mL) for ^{1}H NMR analysis. The samples were vortexed for 10 s and then heated at 50°C for 10 min. This was followed by centrifugation at 13200 rpm for 10 min. The supernatants were pipetted into clean tubes and heated at 90°C for 2 min followed by cooling at 4°C for 30 min and subsequent centrifugation at 13200 rpm for a further 10 min. The supernatants were then transferred to glass vials for UHPLC-MS or 5 mm NMR tubes for NMR analysis. Following the same procedure, 260 mg of milled sample was extracted in 6 mL of H_2O:CH_3OH (4:1 v/v) for fractionation. The resultant extract was aliquoted into a glass autosampler vial for HPLC fractionation.

Fractionation was done on a Dionex UltiMate 3000, Thermo Fisher Scientific HPLC system with an Ascentis C-18 column (5 µm, 5 × 250 mm, Supelco, UK). 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), was used to carry out the chromatographic separation. 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. Following multiple injections (each 100 µL) on the HPLC, the resultant fractions were automatically collected by time into separate glass tubes. The system was set to automatically restart the collection into the same glass tubes at each run. The eluting compounds were monitored between the wavelengths, 200 and 800 nm. After the collection, each fraction (200 µL) was pipetted into a glass vial and analysed by UHPLC-MS. Using a Speedvac concentrator (Genevac, Suffolk, UK), the remaining volume was dried overnight and then dissolved in 700 µL of NMR solvent for subsequent NMR analysis.

4.3.1. 10-hydroxylimonene-7-oic acid-10-O-β-glucopyranoside (**34**)

Yellow amorphous solid (26 µg); UV λ_{max} 220 nm; ^{1}H (600 MHz, D_2O:CD_3OD (4:1)) and ^{13}C NMR (150 MHz, D_2O:CD_3OD (4:1)) data, see Table 1; UHPLC-MS *m/z* 343.1396 $[M-H]^-$ (calcd. For $C_{16}H_{23}O_8^-$, 343.1398).

4.3.2. Ammodauside (**35**)

Yellow amorphous solid (34 µg); UV λ_{max}: 223, 322 nm (**35a**) and 218, 328 nm (**35b**); ^{1}H (600 MHz, D_2O:CD_3OD (4:1)) and ^{13}C NMR (150 MHz, D_2O:CD_3OD (4:1)) data, see Table 2; UHPLC-MS *m/z* 487.1243 $[M-H]^-$ (calcd. for $C_{24}H_{23}O_{11}^-$, 487.1246).

4.3.3. Ammodaucic acid A (**39**)

Amorphous solid (26 µg); UV λ_{max} 223, 307 nm; ^{1}H (600 MHz, D_2O:CD_3OD (4:1)) and ^{13}C NMR (150 MHz, D_2O:CD_3OD (4:1)) data, see Table 3; UHPLC-MS *m/z* 325.0718 $[M-H]^-$ (calcd. For $C_{18}H_{13}O_6^-$, 325.0718).

4.3.4. Ammodaucic acid B (**41**)

Yellow amorphous solid (8 µg); UV λ_{max} 219, 314 nm; ^{1}H (600 MHz, D_2O:CD_3OD (4:1)) and ^{13}C NMR (150 MHz, D_2O:CD_3OD (4:1)) data, see Table 3; UHPLC-MS *m/z* 355.0822 $[M-H]^-$ (calcd. for $C_{19}H_{15}O_7^-$, 355.0823).

4.3.5. Ammodolide (**52**)

Yellow amorphous solid (39 µg); UV λ_{max} 220 nm; ^{1}H (600 MHz, D_2O:CD_3OD (4:1)) and ^{13}C NMR (150 MHz, D_2O:CD_3OD (4:1)) data, see Table 4; UHPLC-MS *m/z* 543.2436 $[M-H]^-$ (calcd. for $C_{26}H_{39}O_{12}^-$, 543.2447).

Funding

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.

CRediT authorship contribution statement

Musa Kisiriko: Writing – original draft, Visualization, Investigation, Formal analysis. **Gabin Bitchagno:** Writing – review & editing, Visualization, Formal analysis. **Clarice Noleto-Dias:** Writing – review & editing, Supervision, Formal analysis. **Imane Naboulsi:** Investigation. **Maria Anastasiadi:** Writing – review & editing, Supervision. **Leon Terry:** Writing – review & editing, Supervision. **Mansour Sobeh:** Writing – review & editing. **Michael Beale:** Writing – review & editing, Supervision, Formal analysis, Conceptualization. **Jane Ward:** Writing – review & editing, Supervision, Project administration, Conceptualization.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Jane Ward reports financial support was provided by OCP Group. Musa

Kisiriko reports financial support was provided by OCP Group. Gabin Bitchagno reports financial support was provided by OCP Group. Clarice Noleto-Dias reports financial support was provided by OCP Group. Maria Anastasiadi reports financial support was provided by OCP Group. Leon Terry reports financial support was provided by OCP Group. Mansour Sobeh reports financial support was provided by OCP Group. Michael Beale reports financial support was provided by OCP Group. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors would like to acknowledge Charlotte Lomax and Gianluca Ruvo for their assistance in the collection of UHPLC-MS and NMR data respectively.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.phytol.2024.11.006.

Data availability

Data will be made available on request.

References

Abderrezag, N., Sánchez-Martínez, J.D., Louaer, O., Meniai, A.-H., Mendiola, J.A., 2021. Optimization of pressurized liquid extraction and in vitro neuroprotective evaluation of *Ammodaucus leucotrichus*. Untargeted metabolomics analysis by UHPLC-MS/MS. Molecules 26, 6951. https://doi.org/10.3390/molecules26226951.

Abderrezag, N., Bragagnolo, F.S., Louaer, O., Meniai, A.-H., Cifuentes, A., Ibáñez, E., Mendiola, J.A., 2024. Optimization of supercritical fluid extraction of bioactive compounds from *Ammodaucus leucotrichus* fruits by using multivariate response surface methodology. J. Supercrit. Fluids 207, 106211. https://doi.org/10.1016/j.supflu.2024.106211.

Annaz, H., Elrashidy, R.A., Mohammed, H.O., Ibrahim, S.E.M., Mahdi, I., Mahmoud, M. F., Sobeh, M., 2024. *Ammodaucus leucotrichus* Coss. & Dur. Attenuates acrylamide-induced hepatic injury through enhancing autophagy and inhibiting endoplasmic reticulum stress and apoptotic signaling. J. Funct. Foods 115, 106090. https://doi.org/10.1016/j.jff.2024.106090.

Belbachir, F.Z., Ramdani, H.L., Decombat, C., Gainche, M., Chalard, P., Djabbou, N., Taibi, N., Bachari, K., Yahia, K.D., Delort, L., Caldefie-Chezet, F., 2024. Chemical profile and assessment of antioxidant, anti-inflammatory, and antiproliferative effects of *Ammodaucus leucotrichus* Coss. & Dur. fruit extracts on human breast cancer. Nat. Prod. Commun. 19 (3). https://doi.org/10.1177/1934578X241239827.

Bellau, M.L.A., Chiurato, M.A., Maietti, A., Fantin, G., Tedeschi, P., Marchetti, N., Tacchini, M., Sacchetti, G., Guerrini, A., 2022. Nutrients and main secondary metabolites characterizing extracts and essential oil from fruits of *Ammodaucus leucotrichus* Coss. & Dur. (Western Sahara). Molecules 27, 5013. https://doi.org/10.3390/molecules27155013.

Bouknana, S., Daoudi, N.E., Bouhrim, M., Ziyyat, A., Legssyer, A., Mekhfi, H., Bnouham, M., 2022. *Ammodaucus leucotrichus* Coss. & Durieu: Antihyperglycemic activity via the inhibition of α-amylase, α-glucosidase, and intestinal glucose absorption activities and its chemical composition. J. Pharm. Pharmacogn. Res. 10, 94–103. https://doi.org/10.56499/jppres21.1141_10.1.94.

Bouknana, S., Lafdil, F.Z., Kandsi, F., Driouech, M., Conte, R., Bouknana, D., Ziyyat, A., Mekhfi, H., Legssyer, A., Bnouham, M., 2024. Antidiabetic and aldose reductase inhibitory activity and LC-MS/MS compositional polyphenol determination of aqueous extract of *Ammodaucus leucotrichus* fruits. Biocatal. Agric. Biotechnol. 58, 103100. https://doi.org/10.1016/j.bcab.2024.103100.

Bouzaid, H., Zinedine, A., Mazzah, A., Msaguad Ayam, I., Ouazzani Chahdi, F., Haoudi, A., Rodi, Y.K., Sfaira, M., Zarrouk, A., Errachidi, F., 2024a. Quantitative and qualitative analysis and evaluation of antioxidant activity of phenolic compounds extracted from apiaceae family spices. Ecol. Eng. Environ. Technol. 25 (4), 45–56. https://doi.org/10.12912/27197050/182946.

Bouzaid, H., Espírito Santo, L., Ferreira, D.M., Machado, S., Costa, A.S.G., Dias, M.I., Calhelha, R.C., Barros, L., Chater, O., Rodi, Y.K., Errachidi, F., Chahdi, F.O., Oliviera, M.B.P.P., Alves, R.C., 2024b. Detailed Phytochemical Composition, Cyto-/Hepatotoxicity, and Antioxidant/Anti-Inflammatory Profile of Moroccan Spices: A Study on Coriander, Caraway, and Mystical Cumin. Molecules 29 (15), 3485. https://doi.org/10.3390/molecules29153485.

Bunzel, M., Ralph, J., Funk, C., Steinhart, H., 2005. Structural elucidation of new ferulic acid-containing phenolic dimers and trimers isolated from maize bran. Tetrahedron Lett. 46, 5845–5850. https://doi.org/10.1016/j.tetlet.2005.06.140.

Dahmane, D., Dob, T., Krimat, S., Nouasri, A., Metidji, H., Ksouri, A., 2017. Chemical composition, antioxidant and antibacterial activities of the essential oils of medicinal plant *Ammodaucus leucotrichus* from Algeria. J. Essent. Oil Res. 29, 48–55. https://doi.org/10.1080/10412905.2016.1201015.

Dall'Acqua, S., Viola, G., Piacente, S., Cappelletti, E.M., Innocenti, G., 2004. Cytotoxic constituents of roots of *Chaerophyllum hirsutum*. J. Nat. Prod. 67, 1588–1590. https://doi.org/10.1021/np040046w.

Dang, T., Zheng, G., Zhang, Q., Jin, P., Zhang, H., Su, L., Qin, D., Yao, G., 2019. Sesquiterpenoids with diverse carbon skeletons from roots of *Cichorium glandulosum* and their anti-inflammatory activities. Fitoterapia 136, 104170. https://doi.org/10.1016/j.fitote.2019.104170.

Es-Safi, I., Mechchate, H., Amaghnouje, A., Calarco, A., Boukhira, S., Noman, O.M., Mothana, R.A., Nasr, F.A., Bekkari, H., Bousta, D., 2020. Defatted hydroethanolic extract of *Ammodaucus leucotrichus* Cosson and Durieu Seeds: Antidiabetic and anti-inflammatory activities. Appl. Sci. 10, 9147. https://doi.org/10.3390/app10249147.

Fan, M., Yuan, S., Li, L., Zheng, J., Zhao, D., Wang, C., Wang, H., Liu, X., Liu, J., 2023. Application of terpenoid compounds in food and pharmaceutical products. Fermentation 9, 119. https://doi.org/10.3390/fermentation9020119.

Hajib, A., Danton, O., Keller, M., Potterat, O., Bougrin, K., Charrouf, Z., Hamburger, M., 2023. Polyacetylenic caffeoyl amides from *Ammodaucus leucotrichus*. Phytochemistry 206, 113555. https://doi.org/10.1016/j.phytochem.2022.113555.

Han, Y.-F., Zhang, Q., Kun, G., Jia, Z.-J., 2005. New Sesquiterpenes from *Sonchus transcaspicus*. Planta Med. 71, 543–547. https://doi.org/10.1055/s-2005-864156.

Hernandez, L.R., Joseph-Nathan, P., 2000. The revised stereostructure of ammolactone. Rev. Lat. De. Quim. 28, 100–104.

Idm'hand, E., Msanda, F., Cherifi, K., 2020. Medicinal uses, phytochemistry and pharmacology of *Ammodaucus leucotrichus*. Clin. Phytosci. 6, 6. https://doi.org/10.1186/s40816-020-0154-7.

Isman, M.B., 2000. Plant essential oils for pest and disease management. Crop Prot. 19, 603–608. https://doi.org/10.1016/S0261-2194(00)00079-X.

Keserla, B., Manivasagam, G.A., Das, A.K., 2023. Isolation and spectral characterisation of bioactives from seaweed *Hydrocharis laevigata (Humb. & Bonpl. ex Willd.) Byng & Christenh.*- using *in silico* & high throughput analytical techniques. Nat. Prod. Res. 1–8. https://doi.org/10.1080/14786419.2023.2283752.

Kim, S.U., Hwang, E.I., Kim, J.R., Jeong, T.S., Lee, S.H., Moon, J.S., Rho, M.-C. 2006. Pharmaceutical composition and health food comprising extract of Phellinus sp. PL3 or phellinsan A isolated from the same as an effective component for prevention and treatment of cardiovascular disease. US2006115492A1, https://worldwide.espacenet.com/patent/search?q=pn%3DUS2006115492A1.

Kurimoto, S.-i, Suzuki, K., Okasaka, M., Kashiwada, Y., Kodzhimatov, O.K., Takaishi, Y., 2012. Sesquiterpene lactone glycosides from the roots of *Ferula varia*. Chem. Pharm. Bull. 60, 913–919. https://doi.org/10.1248/cpb.c12-00350.

Manssouri, M., Znini, M., Majidi, L., 2020. Studies on the antioxidant activity of essential oil and various extracts of *Ammodaucus leucotrichus* Coss. & Dur. Fruits from Morocco. J. Taibah Univ. Sci. 14, 124–130. https://doi.org/10.1080/16583655.2019.1710394.

Masyita, A., Mustika Sari, R., Dwi Astuti, A., Yasir, B., Rahma Rumata, N., Emran, T.B., Nainu, F., Simal-Gandara, J., 2022. Terpenes and terpenoids as main bioactive compounds of essential oils, their roles in human health and potential application as natural food preservatives. Food Chem. X 13, 100217. https://doi.org/10.1016/j.fochx.2022.100217.

Miao, W.G., Nguyen, T., Iqbal, J., Pierens, G.K., Ma, L., Richardson, D.R., Wood, S.A., Mellick, G.D., Quinn, R.J., Feng, Y., 2022. Meeting the challenge 2: Identification of potential chemical probes for Parkinson's disease from *Ligusticum chuanxiong* Hort using cytological profiling. ACS Chem. Neurosci. 13, 2565–2578. https://doi.org/10.1021/acschemneuro.1c00820.

Mouderas, F., Lahfa, F.B., Mezouar, D., Benahmed, N.E.H., 2020. Valorization and identification of bioactive compounds of a spice *Ammodaucus leucotrichus*. Adv. Trad. Med. 20, 159–168. https://doi.org/10.1007/s13596-019-00390-0.

Muckensturm, B., Diyani, F., Nouën, D.L., Fkih-Tetouani, S., Reduron, J.-P., 1997. Ammolactone, a guaianolide from a medicinal plant, *Ammodaucus leucotrichus*. Phytochemistry 44, 907–910. https://doi.org/10.1016/S0031-9422(96)00621-8.

Neghliz-Benabdelkader, H., Benabdelkader, T., Halladj, F., Jullien, F., 2021. Characterization of the chemical diversity in essential oils from vegetative and reproductive organs of *Ammodaucus leucotrichus* subsp. *leucotrichus* Coss. & Dur. growing in Algeria. J. Essent. Oil-Bear. Plants 24, 75–85. https://doi.org/10.1080/0972060X.2020.1871077.

Picmanova, M., Moller, B.L., 2016. Apiose: one of nature's witty games. Glycobiology 26, 430–442. https://doi.org/10.1093/glcob/cww012.

Qiu-An, W., Zhuo-Jia, Z., Chun-Xiang, Z., Hua-Fang, F., 2008. Synthesis of 8-5′ neolignan compounds. Chin. J. Org. Chem. 28, 1605–1610. ⟨https://sioc-journal.cn/Jwk_yjhx/EN/Y2008/V28/I9/1605⟩.

Ralph, J., Quideau, S., Grabber, J.H., Hatfield, R.D., 1994. Identification and synthesis of new ferulic acid dehydrodimers present in grass cell walls. J. Chem. Soc., Perkin Trans. 1 3485–3498. https://doi.org/10.1039/P19940003485.

Rehdorf, J., Kleber, A. 2017. Aqueous compositions of perillic acid compounds. WO2017129339A1, World International Property Organization https://worldwide.espacenet.com/patent/search/family/055275003/publication/WO2017129339A1?q=WO2017129339A1Espacenet – search results.

Šamec, D., Karalija, E., Šola, I., Vujčić Bok, V., Salopek-Sondi, B., 2021. The role of polyphenols in abiotic stress response: The influence of molecular structure. Plants 10, 118. https://doi.org/10.3390/plants10010118.

Selama, G., Cadi, H.E., Ramdan, B., Majdoub, Y.O.E., Dugo, P., Mondello, L., Cacciola, F., Nhiri, M., 2022. Determination of the polyphenolic content of *Ammodaucus leucotrichus* Cosson and Durieu by liquid chromatography coupled with mass

spectrometry and evaluation of the antioxidant and antiglycation properties. J. Sep. Sci. 45, 3301–3309. https://doi.org/10.1002/jssc.202200235.

Toshmatov, Z.O., Li, J., Eshbakova, K.A., Tang, D., Xin, X., Akber Aisa, H., 2019. New monoterpene glucosides from *Dracocephalum komarovi* and their anti-inflammatory activity. Phytochem. Lett. 33, 102–105. https://doi.org/10.1016/j.phytol.2019.08.005.

Wang, Y., Ma, Y.-B., Huang, X.-Y., Li, T.-Z., He, X.-F., Zhang, X.-M., Chen, J.-J., 2023. Artemleucolides A-L, eudesmane-type sesquiterpenoids from *Artemesia leucophylla* and their antihepatoma cytotoxicity. Fitoterapia 165, 105399. https://doi.org/10.1016/j.fitote.2022.105399.

Wu, H.-B., Ma, L.-H., Li, X.-M., Liu, T.T., 2022. Selective phytotoxic effects of sesquiterpenoids from *Sonchus arvensis* as a preliminary approach for the biocontrol of two problematic weeds of wheat. J. Agric. Food Chem. 70, 9412–9420. https://doi.org/10.1021/acs.jafc.2c03462.

Yuan, X., Han, B., Feng, Z.-M., Jiang, J.-S., Yang, Y.-N., Zhang, P.-C., 2020. Chemical constituents of *Ligusticum chuanxiong* and their anti-inflammation and hepatoprotective activities. Bioorg. Chem. 101, 104016. https://doi.org/10.1016/j.bioorg.2020.104016.

Zhang, X., Han, B., Feng, Z.-M., Yang, Y.-N., Jiang, J.-S., Zhang, P.-C., 2018. Ferulic acid derivatives from *Ligusticum chuanxiong*. Fitoterapia 125, 147–154. https://doi.org/10.1016/j.fitote.2018.01.005.

Ziani, B.E.C., Rached, W., Bachari, K., Alves, M.J., Calhelha, R.C., Barros, L., Ferreira, I.C. F.R., 2019. Detailed chemical composition and functional properties of *Ammodaucus leucotrichus* Cross. & Dur. and Moringa oleifera Lamarck. J. Funct. Foods 53, 237–247. https://doi.org/10.1016/j.jff.2018.12.023.