



Københavns Universitet

3-Hydroxy-2-methylpropylglucosinolate and other glucosinolates from *Reseda luteola* (Resedaceae)

De Nicola, Gina R.; Zhu, Jianzhong; Pattison, David I.; Koch, Marcus A.; Staerk, Dan; Agerbirk, Niels

Published in:
Biochemical Systematics and Ecology

DOI:
[10.1016/j.bse.2024.104902](https://doi.org/10.1016/j.bse.2024.104902)

Publication date:
2024

Document version
Publisher's PDF, also known as Version of record

Document license:
[CC BY](#)

Citation for published version (APA):
De Nicola, G. R., Zhu, J., Pattison, D. I., Koch, M. A., Staerk, D., & Agerbirk, N. (2024). 3-Hydroxy-2-methylpropylglucosinolate and other glucosinolates from *Reseda luteola* (Resedaceae). *Biochemical Systematics and Ecology*, 117, Article 104902. <https://doi.org/10.1016/j.bse.2024.104902>



3-Hydroxy-2-methylpropylglucosinolate and other glucosinolates from *Reseda luteola* (Resedaceae)

Gina R. De Nicola^{a,*}, Jianzhong Zhu^b, David I. Pattison^c, Marcus A. Koch^d, Dan Staerk^b, Niels Agerbirk^{c,**}

^a CREA-Research Centre for Vegetable and Ornamental Crops, Via Dei Fiori, 8, 51017, Pescia, Italy

^b Department of Drug Design and Pharmacology, Faculty of Health and Medical Sciences, University of Copenhagen, Universitetsparken 2, DK-2100, Copenhagen, Denmark

^c Department of Plant and Environmental Sciences, University of Copenhagen, Thorvaldsensvej 40, 1871, Frederiksberg C, Denmark

^d Department of Biodiversity and Plant Systematics, Center for Organismal Studies (COS) Heidelberg, Heidelberg University, Im Neuenheimer Feld 345, 69120, Heidelberg, Germany

ARTICLE INFO

Keywords:

Novel glucosinolate
Isomer distinction
Identification confidence levels
Glucosinolate biosynthesis evolution

ABSTRACT

We identified 3-hydroxy-2-methylpropylglucosinolate in foliage of *Reseda luteola* L, although the side chain stereochemistry was not resolved. The glucosinolates (GSLs) in *R. luteola* extracts were bound to an anion exchange column and enzymatically desulfated, followed by isolation by preparative HPLC of the desulfo derivatives, the structures of which were elucidated by NMR and/or MS/MS. Presence of intact 3-hydroxy-2-methylpropylGSL in extracts was supported by direct LC-MS detection and MS2 loss of HSO₄⁻. Prominent GSLs in *R. luteola* foliage were 3-hydroxy-2-methylpropylGSL as well as the previously known indol-3-ylmethylGSL and the 2R (minor) and 2S (major) isomers of 2-hydroxy-2-phenylethylGSL. Minor GSLs included the expected biosynthetic precursors 2-methylpropylGSL (tentative) and phenethylGSL, as well as 1-methylpropylGSL (tentative) and benzylGSL. The 3-hydroxy-2-methylpropylGSL matched a previously reported but so far unidentified constituent in *Limnanthes alba* (Limnanthaceae). The identity of investigated *R. luteola* was confirmed morphologically and from DNA barcoding, and a *Reseda* phylogeny was constructed. In conclusion, 3-hydroxy-2-methylpropylGSL occurs in at least two basal families of the Brassicales order.

1. Introduction

The glucosinolates (GSLs) are amino acid derived secondary plant metabolites, which can be enzymatically converted to bioactive isothiocyanates (Fig. 1) and several other hydrolysis products and downstream biosynthetic products. The GSLs and their products contribute to the taste of many popular vegetables and condiments (Bell et al., 2018; Raffo et al., 2024), and show a varied palette of bioactivities, including agronomic resistance to pests and diseases (Pedras et al., 2011; Poveda et al., 2020; Liu et al., 2021; Plaszkó et al., 2022; Qui et al., 2023; Cárdenas et al., 2023), detrimental effects on domestic animals when fed in large quantities (Di Gioia et al., 2020), and generally beneficial effects of relevant levels in the human nutrition (Marino et al., 2021; Connolly et al., 2021). Importantly, the taste and bioactivity

depend on side chain structure (Bell et al., 2018; Di Gioia and Petropoulos, 2021). The GSLs are biochemically the best understood group of all secondary metabolites, serving as model secondary metabolites (e.g. Sønderby et al., 2010; Kliebenstein and Cacho, 2016; Klein and Sattely, 2017; Edger et al., 2018; Barco and Clay, 2019; Sugiyama et al., 2021; Malhotra et al., 2022; Sanden et al., 2024). For all these reasons, the identification of previously unknown GSLs in nature is of interest, and NMR is a method of choice for GSL structure elucidation (Agerbirk and Olsen, 2012; Ibrahim et al., 2018; Blažević et al., 2020). From plant material dominated by a single GSL, intact GSLs can be isolated and structure elucidated (De Nicola et al., 2014). From plant material that contains a complex mixture of GSLs, structure elucidation is most often based on conversion to desulfoGSLs (dGSLs), which can subsequently be isolated by HPLC using reversed phase columns (Fig. 1). This approach

* Corresponding author.

** Corresponding author.

E-mail addresses: ginarosalinda.denicola@crea.gov.it (G.R. De Nicola), jianzhong.zhu@sund.ku.dk (J. Zhu), dip@plen.ku.dk (D.I. Pattison), marcus.koch@cos.uni-heidelberg.de (M.A. Koch), ds@sund.ku.dk (D. Staerk), nia@plen.ku.dk (N. Agerbirk).

<https://doi.org/10.1016/j.bse.2024.104902>

Received 20 June 2024; Received in revised form 6 September 2024; Accepted 9 September 2024

Available online 16 September 2024

0305-1978/© 2024 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

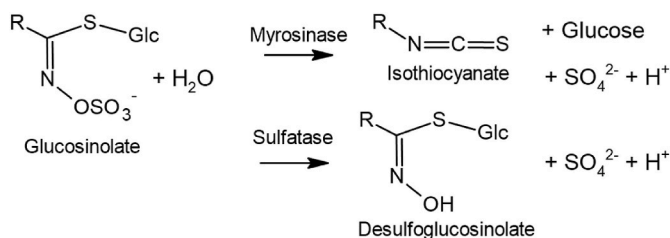


Fig. 1. Hydrolysis reactions of glucosinolates (GSLs). *In vivo*, the native plant enzyme myrosinase may hydrolyze GSLs to isothiocyanates. *In vitro*, sulfatase is used to convert GSLs, either in extracts or adsorbed on anion exchange columns, into easily manageable desulfoGSLs. Both kinds of hydrolysis products have an un-changed side chain (R-group).

was pioneered by Sang et al. (1984) when discovering the 4-substituted indole GSLs, and has been followed by a steady rate of GSL discoveries using the same principle (Blažević et al., 2020; Trabelcy et al., 2021).

In the plant kingdom, GSLs mainly occur in the plant order Brassicales (Blažević et al., 2017; Edger et al., 2018; Barco and Clay, 2019; Agerbirk et al., 2021b) with occasional exceptions (Montaut et al., 2017; Dulović et al., 2024). The core GSL structure (Fig. 1) is considered invariable in plants, except for occasional acylation of the thioglucose moiety, while R-group structures are very diverse (Blažević et al., 2020). The number of natural GSLs known with high certainty (except the stereochemical configuration in some cases) was found to be 88 individual GSLs in a recent review (Blažević et al., 2020). Around a dozen additional “generally acknowledged” structures as well as numerous additional candidates are often discussed (Blažević et al., 2020).

GSL diversity is also of interest in evolutionary biodiversity studies. In such studies, a universal numbering system (Table S1) with a fixed number for each GSL is recommended (Dulović et al., 2023). The list of GSL numbers in a recent review (Blažević et al., 2020) ended with [156] for putative 5-phenylpentylGSL (Dekić et al., 2017). Numbers of proposed but uncertain GSLs, still in need of NMR verification, are written in brackets, such as [156]. Two additional GSL numbers are suggested (Dulović et al., 2023) to represent 3-hydroxy-4-β-D-apiofuranosyloxybenzylGSL (157) (Montaut et al., 2020) and 2-α-L-arabinopyranosyloxybenzylGSL (158) (Trabelcy et al., 2021). In this paper, we report the 91st conclusively elucidated naturally occurring GSL, 3-hydroxy-2-methylpropylGSL (159) (Fig. 2), characterized by a simple structure but an unusual position of the hydroxyl substitution and presence in at least two basal families in the order Brassicales.

Many GSLs are biosynthetically derived from standard amino acids such as Ala, Val, Leu, Ile, Phe, Tyr and Trp, which are converted to corresponding GSLs in what is called the core structure GSL biosynthesis (Sønderby et al., 2010; Harun et al., 2020). In this way, Leu can be converted to 2-methylpropylGSL (62), Phe to benzylGSL (11) and Trp to indol-3-ylmethylGSL (43) (Wang et al., 2020) (Fig. 3). Some amino acids can also undergo chain elongation to higher homologs, an example relevant for this paper is Phe, which can be chain elongated to homoPhe and converted to phenethylGSL (105) (Wang et al., 2021). Those four GSLs all occur in some of the basal families in the Brassicales (Agerbirk et al., 2021b). They are examples of “parent GSLs”, because their side

chain directly reflects the side chain of the precursor amino acid (Fig. 3).

Additional structural diversity is due to various kinds of “decoration” of the parent GSLs, of which hydroxylation by GS-OH enzymes (Fig. 3) is relevant for the present paper (Hansen et al., 2008; Harun et al., 2020; Blažević et al., 2020). The combination of isomeric precursor amino acids and multiple possible kinds and positions of side chain decoration leads to many possible isomeric GSL side chains. Among side chains discussed in this paper, six GSL isomers are known for the side chain C₈H₉O (including 40R and 40S, Fig. 3). For the side chain C₄H₉O, two GSL isomers were previously known with certainty. One was 2-hydroxy-2-methylpropylGSL (31) (Kjær et al., 1956). A second was 1-(hydroxymethyl)propylGSL (30) (Kjær and Christensen, 1962). A third and a fourth isomer, 3-hydroxybutylGSL ([25]) and 4-hydroxybutylGSL ([26]), were suggested from elucidation of structures of natural hydrolysis products but without structure elucidation of the GSLs themselves (Kjær and Schuster, 1971). However, the structure of [26] is backed up by isotope labelling (Graser et al., 2001) and genetic evidence that it is the precursor of a conclusively known GSL (Harun et al., 2020; Blažević et al., 2020). The existence of a fifth isomer (“iso31”) was recently suggested by three of the current authors in a study of the niche oil crop *Limnanthes alba* and another species of the Limnathaceae family, but the structure was not elucidated (Agerbirk et al., 2022). Consequently, we searched for a plant species with levels of “iso31” that were high enough for structure elucidation, and found a candidate in seeds and leaves of *Reseda luteola* L., a traditional plant source of yellow dye. The candidate peak was detected in *R. luteola* seeds and referred to as an “unidentified isomeric hydroxybutylGSL” (Agerbirk et al., 2021a).

The three aims of this paper were (1) to test whether the unidentified hydroxybutylGSL “iso31” in *L. alba* was identical to the one in *R. luteola*, (2) to identify major and minor GSLs in leaves of *R. luteola*, and (3) to test the botanical identity of the source plant and investigate its phylogenetic relationship to previously investigated *Reseda* species.

2. Results and discussion

2.1. Comparison of unidentified glucosinolates in *Limnanthes alba* and *Reseda luteola*

In order to find a suitable source of the apparently undescribed hydroxybutylGSL in *L. alba* (Agerbirk et al., 2022), previously referred to as iso31, candidate plant extracts subjected to enzymatic desulfation of GSLs were compared using UHPLC-MS/MS with stringent mass accuracy. The previously reported but unidentified desulfo derivative in *R. luteola* seed (Agerbirk et al., 2021a), and in *R. luteola* leaves investigated here, matched desulfo iso31 with respect to retention time (*t_R*), accurate *m/z* of [M+H]⁺ and MS2 spectra (Fig. 4). The GSL previously known as iso31, now numbered 159, was also detected in extracts of an independent accession of *R. luteola* (accession UCPH). Hence, the 3-hydroxy-2-methylpropylGSL (159) described here occurred in at least two accessions of *R. luteola* and in at least two species of the family Limnathaceae.

A preliminary quantitative comparison of the dGSL products generated from desulfation using either purified sulfatase or the crude commercially obtained sulfatase was carried out using the UCPH accession. It showed the two sulfatase preparations to be approximately equivalent for *R. luteola* leaf samples under our conditions (Fig. S1). The remaining investigations, described below, were carried out with the main accession from Chiltern Seeds. Preliminary comparison of seeds, roots and leaves suggested that leaves were the best starting material for the structure elucidation of 159.

2.2. Isolation and structure elucidation of desulfated glucosinolates from *Reseda luteola*

Foliage of *R. luteola* was extracted in preparative scale and the GSLs bound to an anion exchange column and converted to dGSLs as for the

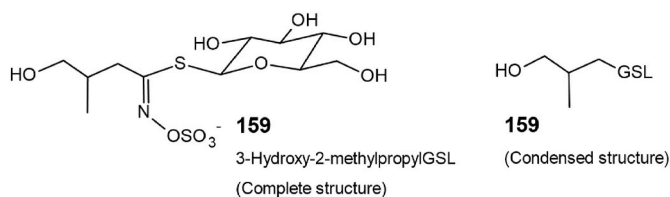


Fig. 2. Structure of 3-hydroxy-2-methylpropylglucosinolate reported here, and the convention for drawing condensed structures with the constant part abbreviated “GSL”.

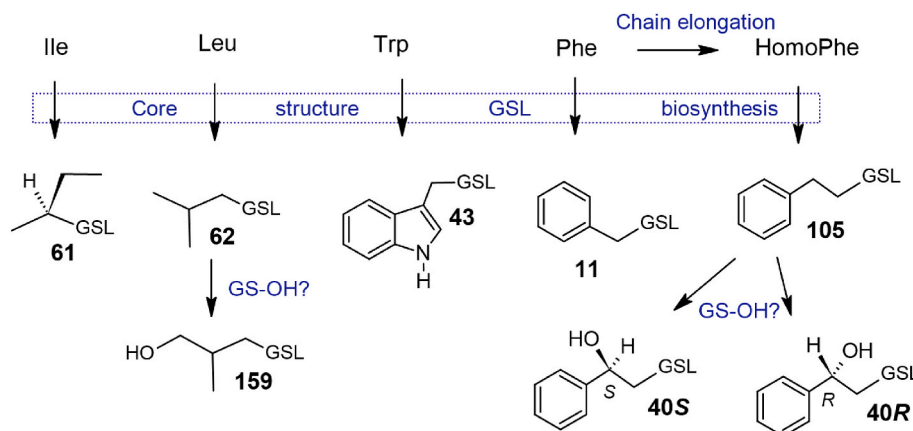


Fig. 3. Glucosinolates demonstrated in *Reseda luteola* and their suggested biosynthetic relationship. The names of the illustrated GSLs are listed in Table S1.

analytical scale extracts. The preparation was concentrated to a small volume and the dGSLs isolated by preparative HPLC, collecting peaks of **d159**, **d40R**, **d40S** and **d43**, as well as four very minor peaks (Fig. S1).

By MS, UV and NMR, **d159** was found to be desulfo 3-hydroxy-2-methylpropylGSL as follows. The dGSL showed an average mass (m/z) of 312.1115 for the $[M+H]^+$ ion in Q-ToF MS ($N = 8$, $SD = 0.0003$), reflecting a molecular formula of $C_{11}H_{21}NO_7S$ (calc. for $C_{11}H_{22}NO_7S^+ = 312.1112$, dev. 1.3 ppm). The MS2 was as expected for an isomeric desulfo hydroxybutylGSL, showing 162 Da and 18 Da losses that could be attributed to loss of anhydroglucose and water from a thioglucoside and a hydroxyl group, respectively (Fig. 4F). The spectral contrast to the MS2 of **d31** suggested **d159** not to be a tertiary alcohol (Fig. 4). The UV spectrum (λ_{max} at ca. 224 nm) agreed with a dGSL (Fig. S1). In 2D NMR, the COSY spectrum suggested two spin systems: $H_2-1 \leftrightarrow H-2 (H_3-4) \leftrightarrow H_2-3$ and $H-1' \leftrightarrow H-2' \leftrightarrow H-3' \leftrightarrow H-4' \leftrightarrow H-5' \leftrightarrow H_2-6'$, reflecting the side chain and the thioglucose moieties, respectively (Fig. S5). The chemical shift of C3 at 65.5 confirmed the side chain to be hydroxylated, and the CH_2 nature of the DEPT peak confirmed a primary alcohol group. HMBC correlations from H-1 and H-1' to $\delta = 153.8$ allowed the identification of C0 as required for the expected dGSL. ROESY correlations between H-1', H-3' and H-5' demonstrated that these protons are on the same side of the ring. The ROESY correlations between H-4' and H_2-6' and weak ROESY between H-1' and H-2' suggested that H-2', H-4', and H-6' are on the other side of the ring, confirming the hexoside to be a glucoside. The chemical shifts of the anomeric proton ($\delta = 4.97$) and carbon ($\delta = 80.6$) confirmed the glucoside to be on S (as opposed to an O), and the coupling constant of the anomeric proton (9.8 Hz) confirmed the β -configuration as required for a dGSL. Key correlation signals supporting the structural elucidation of **d159** are provided in Fig. 5. A detailed summary can be found in Table 1, and all NMR spectra are provided as Figs. S2–S7 (unassigned for teaching purposes in Fig. S8).

2.3. Mass spectral confirmation of intact 3-hydroxy-2-methylpropylglucosinolate

In order to test the hypothesis of the deduced intact **159** being present in the investigated leaves, we searched for the intact **159** using UHPLC-QToF-MS/MS of polar fractions prepared from *R. luteola* leaf extracts ($N = 3$), each analyzed undiluted and 10 fold diluted (Figs. S9–S10). Authentic **40S**, **40R**, **105**, **11** (benzylGSL) and **107** (allylGSL) were analyzed in parallel. The dominating peak in *R. luteola* polar fractions was **40S**, identified by m/z value (found 438.0532 $[M-H]^-$ (mean, $N = 12$, $SD = 0.0007$), calc. for $C_{15}H_{20}O_{10}NS_2^-$: 438.0534, dev. -0.5 ppm), comparison with the authentic reference ($t_R = 10.1$ min) and MS2 fragmentation to previously reported fragments with HSO_4^- at m/z 96.9601 (mean, $N = 12$, $SD = 0.0002$) (calc. for HSO_4^- : 96.9601, dev. -0.4 ppm) being the major fragment as expected for **40S** and other GSLs

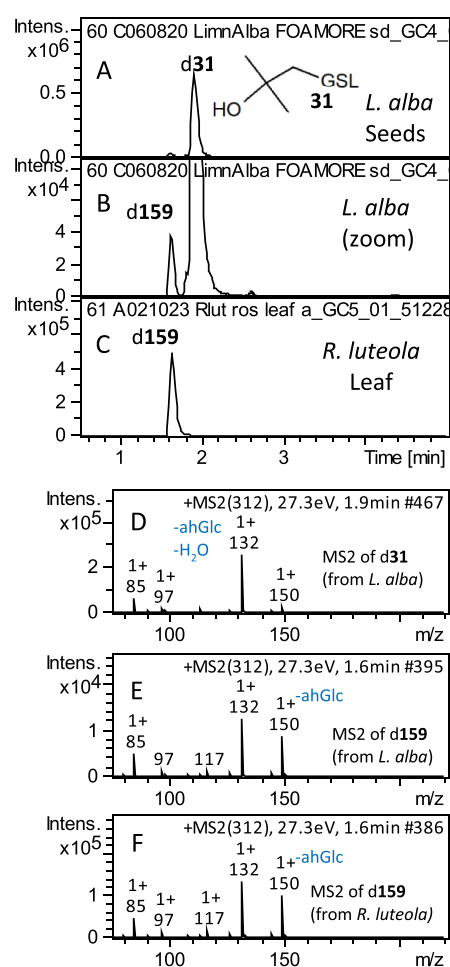


Fig. 4. Distinction and matching of isomeric hydroxybutylglucosinolates by UHPLC retention time (A–C) and MS2 spectra (D–F). A–C: UHPLC-MS chromatograms targeted at desulfo derivatives of **31** and of **159** (m/z 312.1112; 150.0583; 132.0478 ± 0.005) in sulfatase-treated extracts of *Limnanthes alba* (A + B) and *Reseda luteola* (C). Peak detection included proton adducts and ions from loss of anhydroglucose (ahGlc) and water. Desulfo derivatives are indicated by a preceding “d”, i.e., **d31** is desulfo**31**. The **d159** from *R. luteola* coeluted with a minor peak from *L. alba*, suggesting the analytes to be identical. D–F: MS2 spectra of the peaks in panel A–C, as indicated in the individual panels. The MS2 of **d159** from the two species matched (E and F), while the MS2 of **d31** (D) showed a markedly higher tendency for loss of water as expected for a tertiary alcohol, i.e., a relatively low intensity m/z 150 peak, $[M-ahGlc + H]^+$, and high intensity m/z 132 peak, $[M-ahGlc-H_2O + H]^+$.

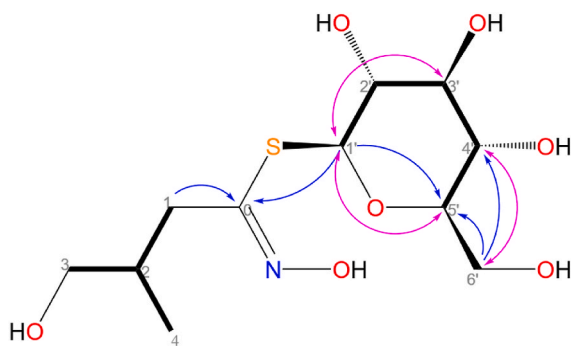


Fig. 5. Molecular connectivity of the side chain to the constant part in desulfo 3-hydroxy-2-methylpropylglucosinolate as revealed by NMR. The bold C-C bonds illustrate COSY correlations and the arrows illustrate selected HMBC (blue single-headed arrows) and ROESY (magenta double-headed arrows) correlations.

(Bianco et al., 2014). A single peak of suitable polarity ($t_R = 1.9$ min) and intensity was observed at the calculated m/z value for the **159** anion (found 390.0534 (mean, $N = 12$, $SD = 0.0012$), calc. for $C_{11}H_{20}O_{10}NS_2^-$: 390.0534, dev. -0.1 ppm). The apparent **159** fragmented in MS2 to the same three known general GSL fragments as observed for **40S**, with HSO_4^- being the major fragment (Fig. 6), all as expected for **159**. The m/z of the HSO_4^- fragment in MS2 from **159** was 96.9606 (mean, $N = 12$, $SD = 0.0007$) (calc. for HSO_4^- : 96.9601, dev. 5.1 ppm). This m/z value was not significantly different from the measured m/z value of HSO_4^- in MS2 from the known **40S** in the same analysis files ($p = 0.18$, Student's t -test, $N = 2 \times 12$). These results were all in agreement with the deduced existence of the intact GSL **159** in *R. luteola*.

The combined evidence for **159** and the derivative **d159** satisfies the consensus requirements for reporting a previously undescribed GSL (Blažević et al., 2020; Trabelcy et al., 2021). The unresolved stereochemistry of the side chain chiral center should be investigated in future studies of hydrolysis products. Based on the well understood GSL biosynthesis using UDP D -glucose (Sønderby et al., 2010 and references cited therein) and extensive synthetic evidence plus x-ray crystallography analysis of model natural GSLs (Blažević et al., 2020 and references cited therein), we illustrate **159** with the glucose being D -glucose and the double bond in (Z) -conformation (Figs. 2 and 4). However, the combined stereochemical uncertainty define the phytochemical

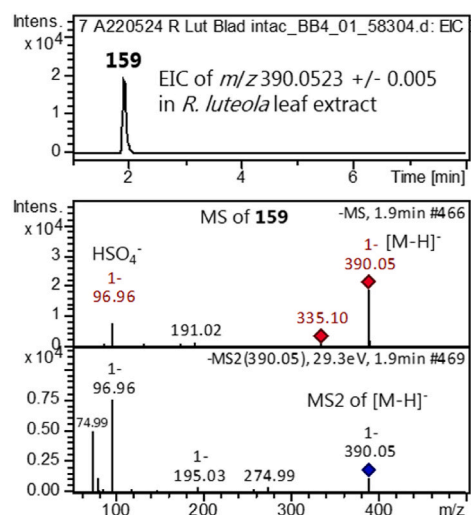


Fig. 6. Chromatogram and mass spectra matching the deduced intact **159** in a polar extract of *Reseda luteola*. The fragments at m/z 274.99 and 74.99 matched previously reported general GSL rearrangement products with sum formulae $C_6H_{11}O_8S_2^-$ and $C_2H_3OS_2^-$, respectively (Bianco et al., 2014). The m/z 195.03 fragment was also observed for the GSLs **11** and **107** at our conditions and is therefore attributed to a general GSL fragment (Fig. S10).

identification confidence level as B (Çiçek et al., 2024).

2.4. NMR evidence for additional glucosinolates in *R. luteola*

Among the additional seven collected peaks (Fig. S1) from preparative HPLC of *R. luteola* dGSLs, three contained sufficient amounts to allow NMR identification. Proton NMR (Figs. S11–S13) and UHPLC-MS/MS of those fractions allowed straightforward identification of **d40R**, **d40S** and **d43**, based on comparison with the known NMR spectra (Agerbirk et al., 2001a, 2001b) and UHPLC-MS/MS match with authentic references (results not shown). This result can be interpreted as identification of **40R**, **40S** and **43** in the investigated *R. luteola* leaves at phytochemical identification confidence level B, adding to the growing list of level B or A-confirmed GSLs in the family Resedaceae (Pagnotta et al., 2020; Trabelcy et al., 2021; Dulović et al., 2023). Peak 8 seemed to represent **d10S** according to UHPLC-MS/MS, while peaks 1, 3

Table 1

^1H and ^{13}C NMR data and 2D COSY, ROESY, and HMBC correlations for desulfo 3-hydroxy-2-methylpropylglucosinolate obtained after desulfation of intact glucosinolates from *Reseda luteola* foliage.

Pos.	δ ^1H , nH, multiplicity (J in Hz)	δ ^{13}C , type	COSY ^a	HMBC	ROESY
1	A: 2.20, 1H, dd (14.7, 9.9 Hz) B: 2.76, 1H, dd (14.9, 4.1 Hz)	34.6, CH ₂	H2 H2, H1A	C0, C2, C3, C4	H1B, H3A, H3B, H1' H3A, H3B, H1'
2	1.99, 1H, m	33.6, CH	H4		H4, H1B, H3A, H3B, H1' H3B
3	A: 3.372, 1H, m B: 3.50, 1H, m	65.5, CH ₂	H2 H2	C1, C2, C4	
4	0.85, 3H, d (6.7 Hz)	14.3, CH ₃		C1, C2, C3	H1A, H3A, H3B
1'	4.97, 1H, d (9.8 Hz)	80.6, CH	H2'	C0, C2', C3', C5'	
2'	3.376, 1H, m	71.5, CH		C1', C3'	H1'
3'	3.494, 1H, m	76.5, CH	H2', H4'	C2', C4'	H1'
4'	3.40, 1H, m	68.5, CH		C5'	H6'A, H6'B
5'	3.48, 1H, m	79.2, CH	H4'		H1', H6'A, H6'B
6'	A: 3.66, 1H, dd (12.6, 5.8 Hz) B: 3.83, 1H, dd (12.5, 2.1 Hz)	59.9, CH ₂	H5' H6'A	C4', C5'	H6'B

Chemical shifts are in ppm. The listed ^{13}C chemical shifts were determined by DEPT-135. In addition, the value for C0 ($\delta = 153.8$) was extracted from HMBC. Multiplicities reported as apparent splittings: d = doublet, dd = doublet of doublets, m = multiplet. "A" denotes the lower chemical shift value and "B" denotes the higher chemical shift value.

^a Reciprocal 2D correlations in COSY and other 2D spectra are indicated only once for each correlation (i.e. the H2' - H3' and H3' - H4' COSY correlations are reported in the H3' caption, but are not repeated in the H2' and H4' captions). The revealing correlations of H2 with all other side chain H is graphically indicated in Fig. S5.

and 4 did not reveal an obvious dGSL component and may have been due to tiny amounts of other types of compounds.

2.5. Additional glucosinolates and analytical scale analysis of *Reseda luteola*

As a control for chemical artifacts during large-scale isolation, analysis of gently treated analytical scale extracts in parallel is recommended (Çiçek et al., 2024). The NMR-identified GSLs were indeed confirmed to be present by UHPLC-MS/MS analysis of analytical scale extracts after sulfatase treatment (Fig. 6; Fig. S14), backed up by authentic references analyzed in parallel Dulović et al. (2024) and results not shown. Estimated levels for four individual leaves were measured, and found to be highly variable (Table 2).

Given the NMR-identified GSLs, we also searched the analytical scale *R. luteola* leaf extracts for potential biosynthetic precursors, based on UHPLC-MS/MS of dGSLs after sulfatase-treatment. The parent GSL from homoPhe, **105**, was thus identified in the leaves and roots, with m/z , t_R and MS/MS matching the authentic reference (phytochemical identification confidence level C) (Çiçek et al., 2024), in agreement with **40R** and **40S** being biosynthesized from **105** as in other species. BenzylGSL (**11**), the parent GSL from Phe, was conclusively detected but was consistently at low levels (Table 2).

Since the newly discovered **159** could be derived from Leu, we searched for and tentatively detected the parent GSL from Leu (**62**), which is the expected biosynthetic precursor of **159**. “Tentative” means that there was match of t_R and m/z with those of authentic references, but MS2 was not obtained due to low level, resulting in phytochemical identification confidence level E (Çiçek et al., 2024). Interestingly, comparable levels of the parent GSL from Ile, 1-methylpropylGSL (**61**), were also tentatively detected, although there was no evidence of a corresponding hydroxyl derivative such as the known 1-(hydroxymethyl)propylGSL (**30**) with a known t_R for the dGSL around 3 min at our conditions (Agerbirk et al., 2022). The apparent presence of **61** and lack of **30** (which would be derived from hydroxylation of **61**) suggested the hydroxylating enzyme converting **62** to **159** to be quite specific. In roots ($N = 3$), levels of **105** were much higher and dominating, in the range 9–35 $\mu\text{mol/g}$ dry wt. (Fig. 7). A number of other GSLs reported from other Resedaceae spp. (Dulović et al., 2023) or considered relevant for other reasons were searched for, supported by authentic references, but not detected (Table 2).

2.6. Phylogeny and verification of the identity of the investigated plant material

As plant species identification has often been somewhat neglected in GSL diversity work, and sometimes explicitly questioned (Kjær and Schuster, 1971; Dulović et al., 2024), we took specific care to obtain

conclusive identification as part of the construction of a *Reseda* phylogeny.

The morphology of bolting and flowering plants was examined by an author (NA) and found to match the description and illustration of *Reseda luteola* L. in a well-established Danish Flora (Frederiksen et al., 2006). Specifically, leaves were entire, lanceolate and with undulating margins, and flowers were pale yellow (Fig. S15).

One representative leaf from the *R. luteola* batch used for isolation of **d159** was split in halves, of which one was subjected to GSL analysis and the other half was subjected to species identification using DNA markers. The GSL profile matched the remaining batch (Table 2, Fig. S16) with a tendency for low levels. The DNA barcoding approach, using the internal transcribed spacer regions 1 and 2 and the 5.8 S rDNA from the nuclear genome and reconstructing the phylogenetic context within the genus *Reseda* (Fig. 8), shows unambiguous clustering with *R. luteola* accessions submitted earlier to Genbank. The structure of the phylogenetic tree (Fig. S17) was identical to an earlier phylogenetic study of genus *Reseda* (Martín-Bravo et al., 2007). In that study the authors investigated 66 out of 85 species from entire Resedaceae family, and *R. luteola* was found very distinct and placed in a single clade A2. The internal and double-control using an additional *R. luteola* collected in Austria (herbarium no. HEID 409163) also grouped with the other *R. luteola* accessions in clade A2.

Both *R. luteola* accessions characterized genetically herein were also studied for plastid DNA variation, which is inherited maternally and often provides additional information on hybridization. For the two markers *trnL* intron/*trnL*F intergenic spacer and the two accessions (GenBank codes: PP885819/PP885820 and PP885821/PP885822) mega-blast searches in GenBank targeted always *R. luteola* accessions with highest matching scores (>99% similarity, 100% query length match) (e.g. matching with DQ987050). As such, we can also exclude any introgression or hybridization with any other *Reseda* species.

2.7. Perspectives

It is generally assumed that 2-hydroxy-2-methylpropylGSL (**31**) (Fig. 4) is biosynthesized by oxidation of the parent GSL from Leu, 2-methylpropylGSL (**62**) (Blažević et al., 2020). Our discovery of the positional isomer **159** suggests that this positional isomer of **31** is likewise biosynthesized from **62**. We find it intriguing that the two positional isomers **31** and **159** are both known from the genus *Reseda*, with **159** in *R. luteola* and **31** in *R. alba* (Dulović et al., 2023). In the genus *Reseda*, we can speculate that a single gene evolved to form either **31** in *R. alba* (Dulović et al., 2023) or **159** in *R. luteola* (Fig. 3). However, both species reside in separate lineages of the genus (Fig. S17) (Martín-Bravo et al., 2007), and more species should be investigated in order to obtain an understanding of GSL profile evolution in this genus (Edger et al., 2018; Barco and Clay, 2019; Agerbirk et al., 2021b). Indeed, the Resedaceae

Table 2

Estimated levels of GSLs (range in $\mu\text{mol/g}$ dry wt.) in individual *R. luteola* leaves and the leaf also used for DNA barcoding, with phytochemical identification confidence levels (ICL) based on the combined spectral and chromatographic information according to Çiçek et al. (2024). Attempted quantification was based on UHPLC-MS after comparison with UHPLC-UV (Section 3.3.).

Amino acid precursor	GSL ^a	Leaves n = 4	DNA-barcode leaf	ICL
Phe	11	0.01–0.1	n.d.	C
HomoPhe	105	0.01–0.7	0.0064	C
HomoPhe	40R	0.3–3	0.96	B
HomoPhe	40S	5–56	3.8	B
Ile	61	0.002–0.02	0.0019	E
Leu	62	0.002–0.07	0.0021	E
Leu	159	0.02–20	0.017	B
Trp	43	0.7–6	0.78	B

^a Names of the GSLs are listed in Table S1. Searched for but not found: **22**, **23**, **[26]**, **30**, **31**, **45**, **46**, **56**, **57**, **109**, **110**, **148**, **157**, **158**. Searched for and not found but absence inconclusive because t_R close to that of the isomer **40R**: **140**. (GSL numbering in general according to Blažević et al. (2020)).

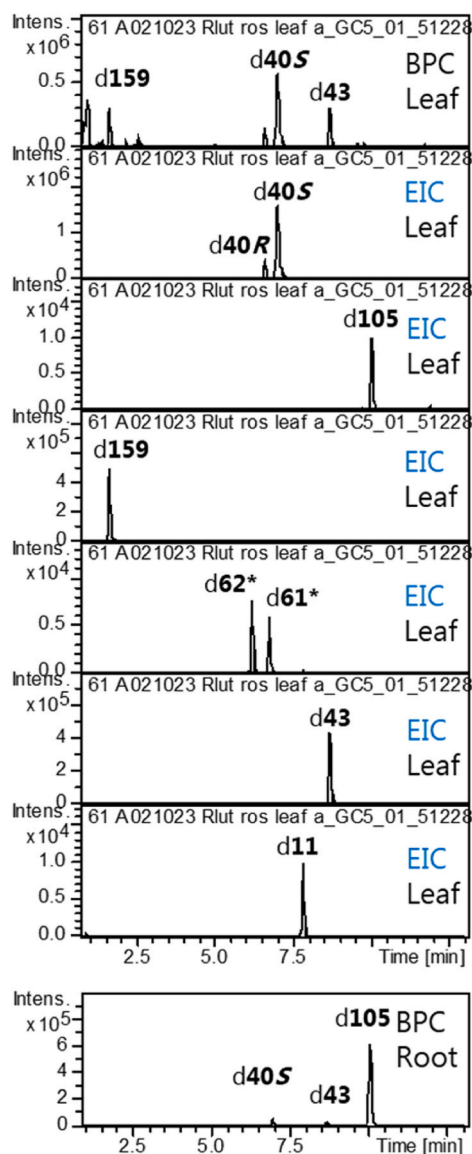


Fig. 7. Chromatographic evidence for the glucosinolate (GSL) profile of *Reseda luteola*. The upper panel is the base peak chromatogram (BPC) from UHPLC-QToF-MS/MS of a desulfated preparation of GSLs from an *R. luteola* leaf, while the lower panel is from a root sample. In the illustrated leaf chromatogram, the Q-ToF detector was moderately overloaded with respect to the dominating d40S in order to obtain good signal to noise ratio for minor peaks, meaning that additional peaks appear more prominent relative to d40S than the true relative abundance. Extracted ion chromatograms (EIC) targeted at individual desulfoGSLs (dGSLs) (Table S1) demonstrated the reported profile of major and minor GSLs in *R. luteola* (Fig. 3). In each case, the peak matched a chromatogram of an authentic standard analyzed in parallel with respect to retention time. Four of the dGSLs (d159, d40R, d40S and d43) were in addition demonstrated by ^1H NMR spectra of the isolated peaks (Section 2.2.). Those four as well as d11 and d105 were likewise confirmed by MS2 spectra matching those of authentic references (Fig. S14), while the identification of d62 and d61 was tentative (indicated by asterisks in the figure) due to lack of MS2 analysis of the low intensity peaks.

family is under continuous investigation and include many more species (Martín-Bravo et al., 2007; Patel et al., 2023) than the relatively few species that have so far been investigated for GSLs (Pagnotta et al., 2020; Dulović et al., 2023). The available phylogeny (Fig. S17) would permit selection of key species for testing for a connection of 31 and 159 in the phylogeny.

Both isomers, 31 and 159, co-occur with 62 in *L. alba* (Fig. 4). Hence,

these positional isomers co-occur in two basal Brassicales families. As GSLs with the hydroxyl group in β -position are very wide-spread in basal and derived families, this decoration is considered ancient (Agerbirk et al., 2021b). It seems possible that the case of *L. alba*, dominated by 31 and with only minor levels of 159, represents an early stage of a transition of the biosynthesis of 31 to the biosynthesis of 159. Possibly, 159 is a minor side product of the reaction forming 31 in *L. alba*. Whatever the mechanism and trajectory, evolutionary changes of the position of GSL side chain hydroxylation seems to have happened in these basal families.

3. Experimental

3.1. General

Water was of milliQ quality, MeCN and MeOH (for LC-MS “HiPerSolv Chromanorm”) were from VWR Chemicals, formic acid (LiChroPur for LC-MS) was from Sigma-Aldrich. AllylGSL (107) was from Karl Roth, Karlsruhe, Germany, benzylGSL (11) was from Merck (Darmstadt, Germany), the mixture of *B. vulgaris* intact GSLs (40S, 40R, 105) was the previously used batch (Müller et al., 2018). Reference dGSLs were the previously reported (Agerbirk et al., 2021a, 2022). In addition, 157 was obtained from S. Montaut (Montaut et al., 2020) and d158 was obtained from B. Trabelcy and Y. Gershman (Trabelcy et al., 2021). UHPLC-QToF-MS/MS was performed using a Dionex Ultimate 3000RS UHPLC system (Thermo Fisher Scientific) coupled to a Compact (QToF) mass spectrometer (Bruker Daltonics) with an electrospray ionization source. The column was a Kinetex XB-C18 UHPLC column 100 m $\times$ 2.1 mm, 1.7 μm , 100 Å pore size (Phenomenex). General operation of the UHPLC-MS instruments and software (DataAnalysis from Bruker Daltonics) has been described elsewhere (Agerbirk et al., 2022). Accurate m/z values were estimated as means of 8–12 individual measurements near the peak apex with 6 decimals from several independently calibrated chromatograms, registered systematically without awareness of the calculated value (Agerbirk et al., 2022). A Student's t-test of m/z for HSO_4^- (two tailed assuming unequal variance) was carried out in MS Excel.

3.2. Plants

Seeds of *Reseda luteola* L. for GSL analysis and NMR identification as well as DNA barcoding and voucher specimen were from Chiltern Seeds (No. 1086 N), Wallingford, UK. Seeds were sown indoors and resulting vigorous rosette plants later cultivated outdoors on loamy soil in the greater Copenhagen area. One representative leaf from two bolting but not yet flowering plants, and root material of one plant, were harvested in July 2023 and dried over silica gel. The remaining leaves from the same two plants were also harvested and allowed to dry on two consecutive warm and dry days, followed by drying and storage over silica gel. A third, co-cultivated plant was allowed to flower and was pressed and deposited as a voucher at Museum Botanicum Hauniense in Copenhagen (CP0014933) (Fig. S15). A second accession of *R. luteola* was material routinely cultivated in the teaching gardens of the Dept. of Plant and Environmental Sciences, University of Copenhagen (Agerbirk et al., 2018). Leaves from this accession were also sampled in 2020 and 2021 and analyzed, and is referred to as “accession UCPH”. The UCPH accession GSL results generally agreed with the Chiltern Seeds accession, but the only result reported here is the comparison of two sulfatase purities (Section 2.1.). A third accession of *R. luteola* was collected in Austria as a botanical control in DNA barcoding (Section 3.7.), and was not used for GSL investigations.

3.3. Analytical glucosinolate extraction and attempted quantification by UHPLC-MS analysis of desulfated derivatives

GSLs were determined by UHPLC-QToF-MS in positive electrospray

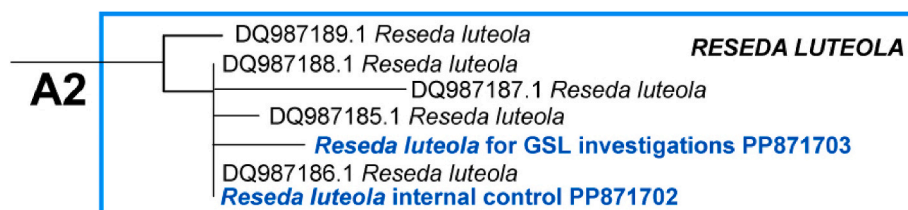


Fig. 8. Detail from phylogenetic reconstruction using ITS and species DNA barcoding confirmation (Fig. S14) of the identity of the investigated leaf material as *Reseda luteola* L (GenBank: PP871703). The internal control from Austria is shown with Genbank code PP871702.

ionization (and parallel UHPLC-UV of major peaks) after desulfation exactly as described previously (Dulović et al., 2024). Briefly, plant parts were dried over silica gel, 0.05 g weighed and extracted for 3×1 min in boiling MeOH/H₂O (7:3) (v:v); extracts were loaded on DEAE-Sephadex anion exchange columns (0.1 g dry wt.) on AcO⁻ form (Agerbirk and Olsen, 2012), which were washed before sulfatase (E.C. 3.1.6.1.) (Sigma-Aldrich) was added and the column left for incubation overnight. Sulfatase was used either as the crude commercial product (0.5 mL at 2.5 g/L) or purified by fractional precipitation according to Graser et al. (2001) or Kiddle et al. (2001), as indicated. For the analyses of individual leaves (Table 2), the crude enzyme was used after it had been found equivalent to the purified. The following day, resulting dGSLs were eluted in 5×1 mL H₂O. Samples were filtered and analyzed by UHPLC-QToF-MS/MS undiluted and 10x diluted in parallel to authentic references and external quantitative reference of benzylGSL (11) (0.5 μmol as std. for UV quantification at 229 nm and 0.05 μmol as std. for attempted MS quantification) based on the total peak area at stringent mass accuracy (± 0.005 Da) from the $[M+H]^+$ ion as well as the prominent fragment ions observed with authentic references (typically ions from loss of anhydroGlc and/or water from the $[M+H]^+$) (Table S1). The UV quantification (assuming equal response factors of 11, 40S and 10S) and MS quantification were roughly equivalent for the dominant d40S and d10S: Levels of 40S based on MS (using the 10x diluted sample because this was in the linear range, and correcting for the dilution), were on average 46% (SD 53%, $N = 7$ leaf and root samples) higher than the levels calculated based on UV (using the undiluted sample, since the UV peaks of diluted samples were diminutive). Similarly, levels of 10S were on average 21% higher based on MS than based on UV (SD 21%, $N = 3$ root samples). MS quantification could not be compared to MS quantification for the other dGSLs since UV peaks were too small. Assuming that the use of all major MS ions would ensure approximately equal ion efficiency, we estimated that MS-based quantification was reliable for semi-accurate quantification also of both high and low intensity peaks, and based all reported levels (Table 2) on MS-quantification.

3.4. Isolation of desulfoglucosinolates for NMR

Dry leaves of *R. luteola* (Chiltern Seeds accession) were reduced to a fine powder. Four 1.0 g samples were extracted for 5 min at 80 °C with 19 mL boiling EtOH/H₂O (7:3) (v:v) using a VDI 25 homogenizer (VWR, Darmstadt, Germany), and then centrifuged with a 5810R centrifuge (Eppendorf AG, Hamburg, Germany) at 4000 rpm for 10 min at 10 °C. The residues were extracted twice as above and all supernatants pooled (48.5 mL). After overnight protein precipitation at -20 °C, the total extract was centrifuged and concentrated under a stream of N₂ to afford 15.0 mL of a bright ochre-yellow clear solution. The crude extract was loaded onto ten DEAE-Sephadex A-25 anion exchange columns (0.2 g dry wt.) conditioned with HPLC grade water (10 mL) (De Nicola et al., 2012). After washing with water (3 mL), sulfatase purified by fractional precipitation according to Kiddle et al. (2001) (300 μL) was loaded onto each column and left overnight at room temperature. Elution of water (4 mL) followed by freeze-drying yielded a crude mixture of dGSLs, which was redissolved in 3 mL ultrapure water. The crude mixture was

separated using an HPLC (LC-40) (Shimadzu, Kyoto, Japan) equipped with an Avantor ACE 5 C18 column (250 × 4.6 mm) (VWR, Darmstadt, Germany) thermostated at 30 °C and having a PDA UV detector. The chromatography was performed injecting 500 μL sample eluting at a flow rate of 1 mL/min with a gradient of water (A) and MeCN (B) following the program: 1 min 1% B; 33 min linear gradient up to 33% B; hold at 33% B for 7 min; 3 min linear gradient down to 1% B. The dGSLs were detected by UV absorbance at 229 nm. Eight peaks (t_R 5.5, 7.5, 13.1, 14.1, 15.0, 15.8, 19.9, 23.0 min) were collected manually right after the UV detector (Fig. S1). Combined fractions from six runs were checked for maximum purity and chromatographed further if needed. Pure fractions were then freeze dried and sent for NMR analysis.

3.5. NMR

NMR spectra were acquired in D₂O on a Bruker Avance III 600 MHz NMR spectrometer (1H operating frequency 600.13 MHz) equipped with a Bruker Sample-Jet autosampler and a 1.7-mm TCI cryoprobe (Bruker Biospin, Karlsruhe, Germany). All NMR data acquisitions were recorded with temperature equilibration to 300 K. Optimization of lock parameters, gradient shimming, and setting of receiver gain were automatically controlled by IconNMR ver. 4.2 (Bruker Biospin, Karlsruhe, Germany). Chemical shifts of ¹H and ¹³C were referenced to tetramethylsilane (δH 0 ppm, δC 0 ppm). ¹H NMR spectra were acquired with 30°-pulses and 64k data points, which were zero-filled to 128k data points. 2D homo- and heteronuclear experiments were acquired with 2048 (DQF-COSY, HMBC, ROESY) or 1730 (multiplicity edited HSQC) data points in the direct dimension and 512 (DQF-COSY, HMBC) or 256 (HSQC, ROESY) data points in the indirect dimension. 2D NMR data were zero-filled to 1k data points in F1 and 4k data points in F2. Processing of NMR data was done using Topspin 4.1.4 (Bruker Biospin, Karlsruhe, Germany).

3.6. Intact GSL extraction and analysis

Polar fractions of *R. luteola* leaves were prepared as follows. For each extract, one dry leaf (ca. 0.03 g), was extracted once for 2 min in 8 mL of boiling usual extraction solvent “70% aq. MeOH”. The extract was dried under a gentle air stream and redissolved in 2 mL 35% aq. MeOH. For clean-up of a polar fraction, a C18 type reversed phase solid phase extraction cartridge with 200 mg column material (Stata C18-E (200 mg), 55 μm, 70A from Phenomenex) was used after being activated with 2 mL of MeOH and equilibrated with 2 mL 35% aq. MeOH. A beaker was placed below the cartridge, and the redissolved extract applied, followed by 3 mL 35% aq. MeOH. Green color stayed on the cartridge. The combined 5 mL eluate was colorless and was filtered through a 0.22 μm filter, and subjected to UHPLC-MS/MS (undiluted as well as diluted 10x with H₂O), along with references (10 μM) of intact 107, 11, and a mixture of 40R, 40S, and 10S. The mobile phase for LC (0.3 mL/min) was composed of (A) 0.05% formic acid (aq) and (B) 0.05% formic acid in MeCN, using the following gradient: 0–2 min, isocratic 98% A with 2% B, the %B was increased linearly to 15% over the next 18 min, then rapidly increased to 90% B over a further 5 min. The column was then washed with 90% B for 5 min, before decreasing back to 2% B over 1

min, and re-equilibrating with 2% B for 6 min before the next injection, resulting in a total method time of 37 min. The column temperature was maintained at 40 °C, samples were cooled at 10 °C, and injection volumes were typically 5 µL. In contrast to the analysis of the desulfated derivatives, for intact GSL analysis the QToF mass spectrometer was run with negative electrospray ionization, with a capillary voltage of 3500 V. All other conditions for the mass spectrometer, including for triggering the data dependent MS/MS acquisition were as described previously (Agerbirk et al., 2021a).

3.7. DNA barcoding and phylogenetic analysis

One leaf of *R. luteola* (Chiltern Seeds accession) from the batch used for GSL extraction was selected and divided in two halves. One half was sent to the Heidelberg lab for DNA barcoding and integrated in HEID collection (dried leaf material HEID8012530), while the other half was subjected to GSL analysis as a control of being representative (Table 2). For DNA barcoding we also studied another proven accession of *R. luteola* collected by MAK in Austria (Lower Austria, Neunkirchen, Gössing, leg. MAK 22.06.2012; HEID409163) as internal control.

For any sample 50 mg dried tissue material was ground in 2 mL Eppendorf tubes using a Retsch swing mill (MM200) in liquid nitrogen and using ceramic and glass beads. DNA extraction was performed with the Invisorb® Spin Plant Mini Kit following the instructions of the supplier (Invitex Diagnostoc, Milton, Australia). DNA was finally eluted with 50 µL elution buffer and DNA concentrations were of about 80 ng/µL (leaf) (Qubit dsDNA HS assay Kit, ThermoFischer Scientific, Karlsruhe, Germany). PCR amplifications of three different DNA markers have been conducted, for which a high number of reference sequences from all over the angiosperms is available via GenBank (<https://www.ncbi.nlm.nih.gov/nucleotide/>), allowing to search similar sequences in GenBank afterwards. From the nuclear genome we selected the internal transcribed spacer regions ITS1 and ITS2 including the intervening 5.8 S rDNA. From the plastid genome we used the highly conserved *trnL* intron and the more variable *trnL-trnF* intergenic spacer (Karl and Koch, 2013). The ITS region was amplified using the primer combination ITS-M13(forward) [TGTAACGACGCGCCAGT-GGAAGGAGAAGTCGT AACAAGG] and ITS-M13(reverse) [CAGGAAACAGCTATGACC-GGGTA ATCCGCGCTGACCTGG]. PCR products spanned the entire ITS1, 5.8 S rDNA, and ITS2 region. The PCR scheme was performed using approximately 25 ng of genomic DNA, 0.5 µL of deoxyribonucleotide triphosphates/dNTPs mix (10 mM), 1 x MangoTaq buffer, 1 µL MgCl₂ (50 mM), 1 U MangoTaq DNA polymerase (Thermo Fischer, Dreieich, Germany) and 0.5 µL of each primer (10 µM), in a total volume of 25 µL. PCR products were purified using the QIAquick PCR Purification Kit Protocol (Qiagen). Amplification was performed using the following protocol: hold at 95 °C for 3 min, 29 cycles of 95 °C for 30 s, 44 °C for 30 s, 72 °C for 1 min, followed 72 °C for 5 min, and a final hold at 4 °C. *TrnL* intron [*trnL*_M13_for: TGTAACGACGCGCCAGT-CGAAATCGGTAGACGCTAC G; *trnL*_M13_rev: CAGGAAACAGCTATGACC-GGGGATAGAGGGACTT GAAC] and *trnLF* [*trnL*_M13_for: TGTAACGACGCGCCAGT-CGAAATC GGTAGACGCTACG; *trnL*_M13_rev: CAGGAAACAGCTATGACC-GGGG ATAGAGGGACTTGAAC]. The PCR scheme followed the protocol: hold at 95 °C for 2 min, 29 cycles of 95 °C for 30 s, 50 °C for 30 s, 72 °C for 1 min, followed 68 °C for 5 min, and a final hold at 4 °C. PCR products were purified using the Wizard SV Gel and PCR Clean-Up System (Promega, Madison, USA) following the manufacturer's instructions. Custom Sanger sequencing was performed at EUROFINS (Ebersberg, Germany) with original PCR products and using standard M13 forward and reverse primers.

For a DNA sequence similarity search the blastn option offered by NCBI was used allowing to run a query sequence against the nucleotide database (<https://www.ncbi.nlm.nih.gov/>). Searches were run against the standard database, restricting the search to Resedaceae and using the megablast (highly similar sequences) option with output size of n = 250. Expected threshold of BLAST hits being expected by chance, with the

observed score or higher was set to 1e-6 (E value).

For the genus *Reseda* and its related genera from the Resedaceae family there is a phylogenetic study available (Martín-Bravo et al., 2007), which applied the same marker genes from the plastid genome, *trnL* intron and *trnLF* intergenic spacer, and the internal transcribed spacers 1 and 2 plus the intervening 5.8S rDNA from the nuclear genome. In this study, phylogenetic analyses placed *R. luteola* distinct from all other species with bootstrap support values of 100%. Thus, we combined our two query sequences with all available *Reseda* ITS1/2 sequences into one single phylogenetic alignment plus various additional *Reseda* accessions archived in GenBank, too. Processing of a final alignment was performed with MUSCLE (Edgar, 2004) and adjusted manually. The TIM3e + G4 model of nucleotide substitution was selected according to ModelFinder (Kalyaanamoorthy et al., 2017) and maximum likelihood (ML) analysis was performed using IQ-Tree (Nguyen et al., 2015) on the IQ-webserver (Trifinopoulos et al., 2016). The number of bootstrap replicates was set to 1000. The alignment is provided with Table S2 and consists of 140 sequences 636 base-pairs in length. In total, 54 different species plus 15 different subspecies and varieties have been considered. As such, the dataset is very comprehensive and covers nearly 65% of all taxonomically accepted species (Martín-Bravo et al., 2007).

CRedit authorship contribution statement

Gina R. De Nicola: Writing – review & editing, Writing – original draft, Methodology, Investigation, Conceptualization. **Jianzhong Zhu:** Writing – review & editing, Writing – original draft, Methodology, Investigation. **David I. Pattison:** Writing – review & editing, Writing – original draft, Methodology, Investigation. **Marcus A. Koch:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Data curation. **Dan Staerk:** Writing – review & editing, Supervision. **Niels Agerbirk:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data are provided as supplementary information. Other data will be available upon request.

Acknowledgements

NA acknowledges R. iori, S. Montaut, B. Trabelcy and Y. Gershman for generous gifts of reference GSLs and dGSLs, Mariela A. G. Ramírez and Jack Olsen for excellent LC-MS support and Torben og Alice Fridomts Fond for financial support for this project. MAK acknowledges excellent lab support from Anna Loreth. The authors thank two anonymous reviewers for constructive comments that improved the final manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.bse.2024.104902>.

References

- Agerbirk, N., Petersen, B.L., Olsen, C.E., Halkier, B.A., Nielsen, J.K., 2001. 1,4-Dimethoxyglucobrassicin in *Barbarea* and 4-hydroxyglucobrassicin in *Arabidopsis* and *Brassica*. J. Agric. Food Chem. 49, 1502–1507. <https://doi.org/10.1021/jf001256r>.

- Agerbirk, N., Olsen, C.E., Nielsen, J.K., 2001b. Seasonal variation in leaf glucosinolates and insect resistance in two types of *Barbarea vulgaris* ssp. *arcuata*. *Phytochemistry* 58, 91–100. [https://doi.org/10.1016/S0031-9422\(01\)00151-0](https://doi.org/10.1016/S0031-9422(01)00151-0).
- Agerbirk, N., Olsen, C.E., 2012. Glucosinolate structures in evolution. *Phytochemistry* 77, 16–45. <https://doi.org/10.1016/j.phytochem.2012.02.005>.
- Agerbirk, N., Matthes, A., Erthmann, P.O., Ugolini, L., Cinti, S., Lazaridi, E., Nuzillard, J.-M., Müller, C., Bak, S., Rollin, P., Lazzari, L., 2018. Glucosinolate turnover in Brassicales species to an oxazolidin-2-one, formed via the 2-thione and without formation of thioamide. *Phytochemistry* 153, 79–83. <https://doi.org/10.1016/j.phytochem.2018.05.006>.
- Agerbirk, N., Hansen, C.C., Olsen, C.E., Kiefer, C., Hauser, T.P., Christensen, S., Jensen, K.R., Ørgaard, M., Pattison, D.I., Lange, C.B.A., Cipollini, D., Koch, M.A., 2021a. Glucosinolate profiles and phylogeny in *Barbarea* compared to other tribe Cardamineae (Brassicaceae) and *Reseda* (Resedaceae), based on a library of ion trap HPLC-MS/MS data of reference desulfo-glucosinolates. *Phytochemistry* 185, 112658. <https://doi.org/10.1016/j.phytochem.2021.112658>.
- Agerbirk, N., Hansen, C.C., Kiefer, C., Hauser, T.P., Ørgaard, M., Lange, C.B.A., Cipollini, D., Koch, M.A., 2021b. Comparison of glucosinolate diversity in the crucifer tribe Cardamineae and the remaining order Brassicales highlights repetitive evolutionary loss and gain of biosynthetic steps. *Phytochemistry* 185, 112668. <https://doi.org/10.1016/j.phytochem.2021.112668>.
- Agerbirk, N., Pattison, D.I., Mandáková, T., Lysak, M., Montaut, S., 2022. Ancient biosynthesis in an oil crop: glucosinolate profiles in *Limnanthes alba* and its relatives (Limnathaceae, Brassicales). *J. Agric. Food Chem.* 70, 1134–1147. <https://doi.org/10.1021/acs.jafc.1c07299>.
- Barco, B., Clay, N.K., 2019. Evolution of glucosinolate diversity via whole-genome duplications, gene rearrangements, and substrate promiscuity. *Annu. Rev. Plant Biol.* 70, 585–604. <https://doi.org/10.1146/annurev-arplant-050718-100152>.
- Bell, L., Oloyede, O.O., Lignou, S., Wagstaff, C., Methven, L., 2018. Taste and flavor perceptions of glucosinolates, isothiocyanates, and related compounds. *Mol. Nutr. Food Res.* 62, e1700990. <https://doi.org/10.1002/mnfr.201700990>.
- Bianco, G., Agerbirk, N., Losito, I., Cataldi, T.R.I., 2014. Acylated glucosinolates with diverse acyl groups investigated by high resolution mass spectrometry and infrared multiphoton dissociation. *Phytochemistry* 100, 92–102. <https://doi.org/10.1016/j.phytochem.2014.01.010>.
- Blažević, I., Montaut, S., Burčul, F., Rollin, P., 2017. Glucosinolates: novel sources and biological potential. In: Mérillon, J.-M., Ramawat, K.G. (Eds.), *Glucosinolates*. Ref. Ser. *Phytochem.*, pp. 3–60. https://doi.org/10.1007/978-3-319-26479-0_1-1.
- Blažević, I., Montaut, S., Burčul, F., Olsen, C.E., Burow, M., Rollin, P., Agerbirk, N., 2020. Glucosinolate structural diversity, identification, chemical synthesis and metabolism in plants. *Phytochemistry* 168, 112100. <https://doi.org/10.1016/j.phytochem.2019.112100>.
- Cárdenas, P.D., Landtved, J.P., Larsen, S.H., Lindegaard, N., Wøhlk, S., Jensen, K.R., Pattison, D.I., Burow, M., Bak, S., Crocoll, C., Agerbirk, N., 2023. Phytoalexins of the crucifer *Barbarea vulgaris*: structural profile and correlation with glucosinolate turnover. *Phytochemistry* 213, 113745. <https://doi.org/10.1016/j.phytochem.2023.113742>.
- Çiçek, S.S., Mangoni, A., Hanschen, F.S., Agerbirk, N., Zidorn, C., 2024. Essentials in the acquisition, interpretation, and reporting of plant metabolite profiles. *Phytochemistry* 220, 114004. <https://doi.org/10.1016/j.phytochem.2024.114004>.
- Connolly, E., Sim, M., Travica, N., Marx, W., Beasy, G., Lynch, G.S., Bondonno, C.P., Lewis, J.S., Hodgson, J.M., Blekkenhorst, L.C., 2021. Glucosinolates from cruciferous vegetables and their potential role in chronic disease: investigating the preclinical and clinical evidence. *Front. Pharmacol.* 12, 767975. <https://doi.org/10.3389/fphar.2021.767975>.
- Dekić, M.S., Radulović, N.S., Stojanović, N.M., Randjelović, P.J., Stojanović-Radić, Z.Z., Najman, S., Stojanović, S., 2017. Spasmolytic, antimicrobial and cytotoxic activities of 5-phenylpentyl isothiocyanate, a new glucosinolate autolysis product from horseradish (*Armoracia rusticana* P. Gaertn., M. Mey. & Scherb., Brassicaceae). *Food Chem.* 232, 329–339. <https://doi.org/10.1016/j.foodchem.2017.03.150>.
- De Nicola, G.R., Nyegue, M., Montaut, S., Iori, R., Menut, C., Tatibouët, A., Rollin, P., Ndoyé, C., Zollo, P.H., 2012. Profile and quantification of glucosinolates in *Pentadiplandra brazzeana* Baillon. *Phytochemistry* 73, 51–56. <https://doi.org/10.1016/j.phytochem.2011.09.006>.
- De Nicola, G.R., Rollin, P., Mazzon, E., Iori, R., 2014. Novel gram-scale production of enantiopure R-sulforaphane from Tuscan black kale seeds. *Molecules* 19, 6975–6986. <https://doi.org/10.3390/molecules19066975>.
- Di Gioia, F., Pinela, J., de Haro, A.B., Ferreira, I.C.F.R., Petropoulos, S.A., 2020. The dilemma of “good” and “bad” glucosinolates and the potential to regulate their content. In: Galanakis, C.M. (Ed.), *Glucosinolates: Properties, Recovery, and Applications*. Academic Press, Cambridge, Massachusetts, pp. 1–45. <https://doi.org/10.1016/B978-0-12-816493-8.00001-9>.
- Di Gioia, F., Petropoulos, S.A., 2021. Glucosinolates. In: Galanakis, C.M. (Ed.), *Food Bioactives and Health*. Springer, Cham. https://doi.org/10.1007/978-3-030-57469-7_2.
- Dulović, A., Tomaš, J., Blažević, I., 2023. Glucosinolates in wild-growing *Reseda* spp. from Croatia. *Molecules* 28, 1753. <https://doi.org/10.3390/molecules28041753>.
- Dulović, A., Koch, M.A., Thongyoo, P., Pattison, D.I., Blažević, I., Rollin, P., Agerbirk, N., 2024. Glucosinolates in non-Brassicaceae plant species: critical literature evaluation and testing of two high chemical quality reports. *Biochem. Systemat. Ecol.* 116, 104864. <https://doi.org/10.1016/j.bse.2024.104864>.
- Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. <https://doi.org/10.1093/nar/gkh340>.
- Edger, P.P., Hall, J.C., Harkess, A., Tang, M., Coombs, J., Mohammadin, S., Schranz, M. E., Xiong, Z., Leebens-Mack, J., Meyers, B.C., Sysma, K.J., Koch, M.A., Al-Shehbaz, I. A., Pires, J.C., 2018. Brassicales phylogeny inferred from 72 plastid genes: a reanalysis of the phylogenetic localization of two paleopolyploid events and origin of novel chemical defenses. *Am. J. Bot.* 105, 463–469. <https://doi.org/10.1002/ajb2.1040>.
- Frederiksen, S., Rasmussen, F.N., Seberg, O., 2006. *Dansk Flora*. Gyldendal, København, p. 701.
- Graser, G., Oldham, N.J., Brown, P.D., Temp, U., Gershenzon, J., 2001. The biosynthesis of benzoic acid glucosinolate esters in *Arabidopsis thaliana*. *Phytochemistry* 57, 23–32. [https://doi.org/10.1016/S0031-9422\(00\)00501-X](https://doi.org/10.1016/S0031-9422(00)00501-X).
- Hansen, B.G., Kerwin, R.E., Ober, J.A., Lambrix, V.M., Mitchell-Olds, T., Gershenzon, J., Halkier, B.A., Kliebenstein, D.J., 2008. A novel 2-oxoacid-dependent dioxygenase involved in the formation of the goiterogenic 2-hydroxybut-3-enyl glucosinolate and generalist insect resistance in *Arabidopsis*. *Plant Physiol.* 148, 2096–2108. <https://doi.org/10.1104/pp.108.129981>.
- Harun, S., Rehda, Abdullah-Zawawi, M.-R., Goh, H.-H., Mohamed-Hussein, Z.-A., 2020. A comprehensive gene inventory for glucosinolate biosynthetic pathway in *Arabidopsis thaliana*. *J. Agric. Food Chem.* 68, 7281–7297. <https://doi.org/10.1104/pp.18.00998>.
- Ibrahim, N., Allart-Simon, I., De Nicola, G.R., Iori, R., Renault, J.-H., Rollin, P., Nuzillard, J.-M., 2018. Advanced NMR-based structural investigation of glucosinolates and desulfo-glucosinolates. *J. Nat. Prod.* 81, 323–334. <https://doi.org/10.1021/acs.jnatprod.7b00776>.
- Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., von Haeseler, A., Jermin, L.S., 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* 14, 587–589. <https://doi.org/10.1038/nmeth.4285>.
- Karl, R., Koch, M.A., 2013. A world-wide perspective on crucifer speciation and evolution: phylogeny, biogeography and trait evolution in tribe Arabideae. *Ann. Bot.* 112, 983–1001. <https://doi.org/10.1093/aob/mct165>.
- Kiddle, G., Bennett, R.N., Botting, N.P., Davidson, N.E., Robertson, A.A., Wallsgrove, R. M., 2001. High-performance liquid chromatographic separation of natural and synthetic desulphoglucosinolates and their chemical validation by UV, NMR and chemical ionisation-MS methods. *Phytochem. Anal.* 12, 226–242. <https://doi.org/10.1002/pca.589>.
- Klein, A.P., Sattely, E.S., 2017. Biosynthesis of cabbage phytoalexins from indole glucosinolates. *Proc. Natl. Acad. Sci. U.S.A.* 114, 1910–1915. <https://doi.org/10.1073/pnas.1615625114>.
- Kliebenstein, D.J., Cacho, N.I., 2016. Nonlinear selection and a blend of convergent, divergent and parallel evolution shapes natural variation in glucosinolates. *Adv. Bot. Res.* 80, 31–55. <https://doi.org/10.1016/bs.abr.2016.06.002>.
- Kjær, A., Gmelin, R., Jensen, R.B., 1956. Isothiocyanates. XVI. Glucosinonin, the natural precursor of 5,5-dimethyl-2-oxazolidinethione. *Acta Chem. Scand.* 10, 432–438. <https://doi.org/10.3891/acta.chem.scand.10-0432>.
- Kjær, A., Christensen, B.W., 1962. Isothiocyanates. XL. Glucosinistrin, a novel glucoside present in seeds of *Sisymbrium austriacum* Jacq. *Acta Chem. Scand.* 16, 71–77. <https://doi.org/10.3891/acta.chem.scand.16-0071>.
- Kjær, A., Schuster, A., 1971. Glucosinolates in *Capparis flexuosa* of Jamaican origin. *Phytochemistry* 10, 3155–3160. [https://doi.org/10.1016/S0031-9422\(00\)97368-0](https://doi.org/10.1016/S0031-9422(00)97368-0).
- Liu, Z., Wang, H., Xie, J., Lv, J., Zhang, G., Hu, L., Luo, S., Li, L., Yu, J., 2021. The roles of Cruciferae glucosinolates in disease and pest resistance. *Plants* 10, 1097. <https://doi.org/10.3390/plants10061097>.
- Malhotra, B., Kumar, P., Bisht, N., 2022. Defense versus growth trade-offs: insights from glucosinolates and their catabolites. *Plant Cell Environ.* 2022, 1–21. <https://doi.org/10.1111/pce.14462>.
- Marino, M., Martini, D., Venturi, S., Tucci, M., Porrini, M., Riso, P., Del Bo, C., 2021. An overview of registered clinical trials on glucosinolates and human health: the current situation. *Front. Nutr.* 8, 730906. <https://doi.org/10.3389/fnut.2021.730906>.
- Martín-Bravo, S., Meimberg, H., Luceno, M., Märkl, W., Valcárcel, V., Bräucher, C., Vargas, P., Heubl, G., 2007. Molecular systematics and biogeography of Resedaceae based on ITS and trnL-F sequences. *Mol. Phylogenet. Evol.* 44, 1105–1120. <https://doi.org/10.1016/j.ympev.2006.12.016>.
- Montaut, S., De Nicola, G.R., Agnani, H., Issembe, Y., Rollin, P., Menut, C., 2017. Probing for the presence of glucosinolates in three *Drypetes* spp. (*Drypetes euryodes* (hiern) Hutch., *Drypetes gossweileri* S. Moore, *Drypetes lacinata* Hutch.) and two *Rinorea* spp. (*Rinorea subintegrifolia* O. Ktze and *Rinorea woermanniana* (büttner) Engl.) from Gabon. *Nat. Prod. Res.* 31, 308–313. <https://doi.org/10.1080/14786419.2016.1236099>.
- Montaut, S., Read, S., Blažević, I., Nuzillard, J.-M., Roje, M., Harakat, D., Rollin, P., 2020. Investigation of the glucosinolates in *Hesperis matronalis* L. and *Hesperis lacinata* All.: Unveiling 4'-O-β-D-apiofuranosylglucosinonitrin. *Carbohydr. Res.* 488, 107898. <https://doi.org/10.1016/j.carres.2019.107898>.
- Müller, C., Schulz, M., Pagnotta, E., Ugolini, L., Yang, T., Matthes, A., Lazzari, L., Agerbirk, N., 2018. The role of the glucosinolate-myrosinase system in mediating greater resistance of *Barbarea verna* than *B. vulgaris* to *Mamestra brassicae* larvae. *J. Chem. Ecol.* 44, 1190–1205. <https://doi.org/10.1007/s10886-018-1016-3>.
- Nguyen, L.-T., Schmidt, H.A., von Haeseler, A., Minh, B.Q., 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum likelihood phylogenies. *Mol. Biol. Evol.* 32, 268–274. <https://doi.org/10.1093/molbev/msu300>.
- Pagnotta, E., Montaut, S., Matteo, R., Rollin, P., Nuzillard, J.M., Lazzari, L., Bagatta, M., 2020. Glucosinolates in *Reseda lutea* L.: distribution in plant tissues during flowering time. *Biochem. Systemat. Ecol.* 90, 104043. <https://doi.org/10.1016/j.bse.2020.104043>.
- Patel, R.M., Prajapati, S.R., Martín-Bravo, S., 2023. *Ochradenus lakhpatensis* (Resedaceae), a new species from Gujarat, India. *Phytotaxa* 633, 92–100. <https://doi.org/10.11646/phytotaxa.633.1.10>.

- Pedras, M.S.C., Yaya, E.E., Glawischig, E., 2011. The phytoalexins from cultivated and wild crucifers: chemistry and biology. *Nat. Prod. Rep.* 28, 1381–1405. <https://doi.org/10.1039/c1np00020a>.
- Plaszkó, T., Szűcs, Z., Vasas, G., Gonda, S., 2022. Interactions of fungi with non-isothiocyanate products of the plant glucosinolate pathway: a review on product formation, antifungal activity, mode of action and biotransformation. *Phytochemistry* 200, 113245. <https://doi.org/10.1016/j.phytochem.2022.113245>.
- Poveda, J., Francisco, M., Carrea, M.E., Velasco, P., 2020. Development of transgenic *Brassica* crops against biotic stresses caused by pathogens and arthropod pests. *Plants* 9, 1664. <https://doi.org/10.3390/plants9121664>.
- Qui, H., King, G., Borpatragohain, P., Zou, J., 2023. Developing multifunctional crops by engineering Brassicaceae glucosinolate pathways. *Plant Comm* 4, 100565. <https://doi.org/10.1016/j.xplc.2023.100565>.
- Raffo, A., Baiamonte, I., De Nicola, G.R., Melini, V., Moneta, E., Nardo, N., Peparao, M., Saggia Civitelli, E., Sinesio, F., 2024. Sensory attributes driving preference for wild rocket (*Diplotaxis tenuifolia*) leaves tasted as a single ingredient and as a part of a recipe. *Foods* 13, 1699. <https://doi.org/10.3390/foods13111699>.
- Sanden, N.C.H., Kanstrup, C., Crocoll, C., Schulz, A., Nour-Eldin, H.H., Halkier, B.A., Xu, D., 2024. An UMAMIT-GTR transporter cascade controls glucosinolate seed loading in *Arabidopsis*. *Nat. Plants* 10, 172–179. <https://doi.org/10.1038/s41477-023-01598-4>.
- Sang, J.P., Minchinton, I.R., Johnstone, P.K., Truscott, R.J.W., 1984. Glucosinolate profiles in the seed, root and leaf tissue of cabbage, mustard, rapeseed, radish and swede. *Can. J. Plant Sci.* 64, 77–93. <https://doi.org/10.4141/cjps84-011>.
- Sugiyama, R., Li, R., Kuwahara, A., Nakabayashi, R., Sotta, N., Mori, T., Ito, T., Ohkama-Ohtsu, N., Fujiwara, T., Saito, K., Nakano, R.T., Bednarek, P., Hirai, M.Y., 2021. Retrograde sulfur flow from glucosinolates to cysteine in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA* 118, e2017890118.
- Sønderby, I.E., Geu-Flores, F., Halkier, B.A., 2010. Biosynthesis of glucosinolates – gene discovery and beyond. *Trends Plant Sci.* 15, 283–290. <https://doi.org/10.1016/j.tplants.2010.02.005>.
- Trabelcy, B., Chinkov, N., Samuni-Blank, M., Merav, M., Izhaki, I., Carmeli, S., Gerchman, Y., 2021. Investigation of glucosinolates in the desert plant *Ochradenus baccatus* (Brassicales: Resedaceae). Unveiling glucochradenin, a new arabinosylated glucosinolate. *Phytochemistry* 187, 112760. <https://doi.org/10.1016/j.phytochem.2021.112760>.
- Trifinopoulos, J., Nguyen, L.-T., von Haeseler, A., Minh, B.Q., 2016. W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Res.* 44 (W1), W232–W235. <https://doi.org/10.1093/nar/gkw256>.
- Wang, C., Dissing, M.M., Agerbirk, N., Crocoll, C., Halkier, B.A., 2020. Characterization of *Arabidopsis* CYP79C1 and CYP79C2 by glucosinolate pathway engineering in *Nicotiana benthamiana* shows substrate specificity toward a range of aliphatic and aromatic amino acids. *Front. Plant Sci.* 11, 57. <https://doi.org/10.3389/fpls.2020.00057>.
- Wang, C., Dissing, M.M., Agerbirk, N., Crocoll, C., Halkier, B.A., 2021. Engineering and optimization of the 2-phenylethylglucosinolate production in *Nicotiana benthamiana* by combining biosynthetic genes from *Barbarea vulgaris* and *Arabidopsis thaliana*. *Plant J.* 106, 978–992. <https://doi.org/10.1111/tpj.15212>.