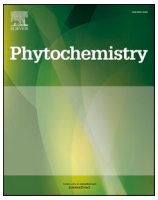
Phytochemistry 185 (2021) 112658

Contents lists available at ScienceDirect

Phytochemistry

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



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 desulfoglucosinolates

Niels Agerbirk [a,*], Cecilie Cetti Hansen [a], Carl Erik Olsen [a], Christiane Kiefer [b], Thure P. Hauser [a], Stina Christensen [a], Karen R. Jensen [a], Marian Ørgaard [a], David I. Pattison [a], Conny Bruun Asmussen Lange [a], Don Cipollini [c], Marcus A. Koch [b]

[a] Department of Plant and Environmental Sciences, University of Copenhagen, Thorvaldsensvej 40, 1871, Frederiksberg C, Denmark
[b] Department of Biodiversity and Plant Systematics, Centre for Organismal Studies, Heidelberg University, 69120, Heidelberg, Germany
[c] Department of Biological Sciences, Wright State University, 3640 Colonel Glenn Highway, Dayton, OH, 45435, USA

ARTICLE INFO

Keywords:
Arabidopsis
Armoracia
Barbarea
Nasturtium
Planodes
Rorippa
Brassicaceae
Reseda
Resedaceae
Glucosinolate identification
Mass spectrometry
Diversity
Polymorphism
Phylogeny

ABSTRACT

A library of ion trap MS2 spectra and HPLC retention times reported here allowed distinction in plants of at least 70 known glucosinolates (GSLs) and some additional proposed GSLs. We determined GSL profiles of selected members of the tribe Cardamineae (Brassicaceae) as well as *Reseda* (Resedaceae) used as outgroup in evolutionary studies. We included several accessions of each species and a range of organs, and paid attention to minor peaks and GSLs not detected. In this way, we obtained GSL profiles of *Barbarea australis*, *Barbarea grayi*, *Planodes virginica* selected for its apparent intermediacy between *Barbarea* and the remaining tribe and family, and *Rorippa sylvestris* and *Nasturtium officinale*, for which the presence of acyl derivatives of GSLs was previously untested. We also screened *Armoracia rusticana*, with a remarkably diverse GSL profile, the emerging model species *Cardamine hirsuta*, for which we discovered a GSL polymorphism, and *Reseda luteola* and *Reseda odorata*. The potential for aliphatic GSL biosynthesis in *Barbarea vulgaris* was of interest, and we subjected P-type and G-type *B. vulgaris* to several induction regimes in an attempt to induce aliphatic GSL. However, aliphatic GSLs were not detected in any of the *B. vulgaris* types. We characterized the investigated chemotypes phylogenetically, based on nuclear rDNA internal transcribed spacer (ITS) sequences, in order to understand their relation to the species *B. vulgaris* in general, and found them to be representative of the species as it occurs in Europe, as far as documented in available ITS-sequence repositories. In short, we provide GSL profiles of a wide variety of tribe Cardamineae plants and conclude aliphatic GSLs to be absent or below our limit of detection in two major evolutionary lines of *B. vulgaris*. Concerning analytical chemistry, we conclude that availability of authentic reference compounds or reference materials is critical for reliable GSL analysis and characterize two publicly available reference materials: seeds of *P. virginica* and *N. officinale*.

1. Introduction

Glucosinolates (GSLs) are secondary metabolites with a fascinating chemistry (Blažević et al., 2020) and biology (e.g. Sun et al., 2019) and the subject of much interest in evolutionary biology (e.g., Edger et al., 2015). The GSLs comprise a moderate number of structures, GSL structure elucidation is relatively simple (Blažević et al., 2020), and GSL biosynthesis is increasingly well-understood (Sønderby et al., 2010), making the GSLs an obvious target for detailed evolutionary investigation.

There are three experimental requirements for investigating the evolution of any class of metabolite: conclusive species identification, availability of reliable phylogenies, and conclusive chemical analysis. The first requirement has been possible for many years but is still a common problem in phytochemical literature (Zidorn, 2017). Reliable molecular phylogenies of plants are a recent development, which has steadily improved but has not yet reached perfection (e.g., Edger et al., 2018). For GSLs, reliable routine analysis became possible with liquid

* Corresponding author.
E-mail address: nia@plen.ku.dk (N. Agerbirk).

https://doi.org/10.1016/j.phytochem.2021.112658
Received 5 May 2020; Received in revised form 30 December 2020; Accepted 3 January 2021

chromatography hyphenated to MS and development of suitable protocols backed up by authentic references. Characterization of the GSL profile in *Arabidopsis thaliana* (L.) Heynh. (e.g. Reichelt et al., 2002; Brown et al., 2003) set the quality standard needed for GSL profiling of any species. That quality standard is realistic and can be summarized by the need for HPLC-MS/MS supported by authentic references and NMR elucidation of any remaining structures (Blažević et al., 2020). For a limited set of GSLs, GC-MS of isothiocyanates after controlled hydrolysis by added myrosinase is equivalent when backed up by authentic standards (Daxenbichler et al., 1991).

It is only in recent years that reliable investigation of GSL diversity in a reliable phylogenetic perspective has been realized. Classical research reached meaningful conclusions for a subset of structures at the family and genus levels (Kjær, 1976; Daxenbichler et al., 1991). An extensive update (Fahey et al., 2001) is unreliable in many aspects (Agerbirk and Olsen, 2012; Blažević et al., 2020) and furthermore outdated. The modern pioneering paper in the field was by Windsor et al. (2005). Only a few scientific papers have since presented structurally conclusive GSL profiles of a range of species in a phylogenetic perspective. One was a primary paper (Agerbirk et al., 2008) and a second was a primary paper backed up by critical literature review (Olsen et al., 2016). A recent primary paper (Czerniawski et al., 2021) used *A. thaliana* seeds as reference material and reliably scored related species for GSLs in multiple organs (except conclusive evidence for suggested existence of the uncertain "*n*-hexylGSL" and "5-oxooctylGSL" in *A. thaliana* were not reported). Finally, a critical review discussed reliable reports of GSL diversity in a phylogenetic perspective, including detailed attention to non-Brassicaceae families (Blažević et al., 2017).

Other recent high-quality reports attempting to compare GSL diversity with phylogenetic trees are explicitly inconclusive concerning GSL identities but have nonetheless demonstrated the potential of such studies (Mithen et al., 2010; Cacho et al., 2015; Rapo et al., 2019). In essence, HPLC-MS without comparison with authentic standards is too equivocal for concluding GSL identities. A good example is suggested GSL identification by HPLC-MS that was largely not supported by use of authentic standards (Rapo et al., 2019). This report has recently been corrected by the authors due to multiple possible isomers (correction March 17, 2020 with the electronic version of the paper). Likewise, a recent (impressive) combined genomics-transcriptomics-metabolomics investigation of the genus *Erysimum* (Züst et al., 2020) could not distinguish positions of hydroxylation and specific isomers of possibly branched-side chain GSLs using MS-evidence only. Two of the listed 25 GSL names were explicitly inconclusive, two more were neither among conclusively nor tentatively known GSLs listed in general reviews, and the bases for distinction of isomeric butylGSLs were not specified. It is obvious that there are general obstacles for reliable GSL identification in the current literature, even in (otherwise) state-of-the-art papers. Identification problems are not limited to evolutionary GSL reports, and four requirements for reliable identification and transparent reporting of GSL profiles using HPLC-MS were recently concluded (Blažević et al., 2020).

Many investigations of the evolution of general GSL biosynthesis have focused on gene sequence evolution, as detailed in an accompanying paper, and rely, either explicitly or implicitly, on tabulated data of GSL diversity of considerable age and moderate reliability. The very crude GSL characters investigated so far (e.g., Edger et al., 2018) are basically use of either Trp, chain elongated Met or other amino acids as biosynthetic precursors based on rather old GSL diversity tables. However, to understand in structural detail the evolution and current diversity and ecology in the most biochemically complex and species rich Brassicales family, Brassicaceae, requires critically assessed, fully reliable GSL profiles.

A key missing resource for improved GSL diversity data is a sufficiently wide selection of reference compounds or materials, allowing isomer distinction and conclusive search for critical GSLs. Systematic screening of reliably identified plant species under well-calibrated, non-targeted analytical conditions backed up by references, is needed. We have elsewhere described our attempts at obtaining such data (Olsen et al., 2016). The present report follows up on specific identified problem areas in that paper, including a less than complete range of reference compounds used in 2016 and specific uncertainties in the literature on some species.

We have concentrated on members of the tribe Cardamineae as an evolutionary study system, because the tribe is of general ecological and even agronomical interest. An extensive molecular phylogeny of this tribe is reported in an accompanying paper (Agerbirk et al., 2021). The tribe is sufficiently defined phylogenetically and contains several actively investigated species. These include the crops watercress, *Nasturtium officinale* W.T. Aiton (Agerbirk et al., 2014; Voutsina et al., 2016; Jeon et al., 2017; Pedras and To, 2018). and horseradish, *Armoracia rusticana* P. Gaertn., B. Mey & Sherb (Agneta et al., 2014; Dekić et al., 2017; Mandákova and Lysak, 2019). Additional species include *Cardamine pratensis* L., investigated for its variable, atypical GSL profile (Agerbirk et al., 2010) and cytological evolution in a species complex (Melichárková et al., 2020) and *Cardamine diphylla* (Michx.) O. Schwarz investigated from a traditional medicine point of view (Montaut and Bleeker, 2013). In particular, the tribe includes the emerging general model species *Cardamine hirsuta* L. (for references, see the accompanying paper) and the much studied, multi-resistant species *Barbarea vulgaris* W.T. Aiton. For the crops in the tribe, we use common names in the rest of the text, while we use scientific names for the remaining species. Nomenclature within Brassicaceae follows https://brassibase.cos.uni-heidelberg.de, and the remaining follow www.worldfloraonline.org and www.ipni.org.

The genus *Barbarea* is well-supported as monophyletic, although relationships among species remain to be resolved (Lange et al., in review), and deviates chemically from other members of the family (Agerbirk et al., 2003a; Byrne et al., 2017). The species *B. vulgaris* in particular has been well studied for its resistance to a broad range of insects and pathogens (see below), as a "dead-end" trap crop (Badenes-Perez et al., 2017), for saponin biosynthesis (Erthmann et al., 2019), for phytoalexin biosynthesis (Pedras et al., 2015; Pedras and To, 2018), for GSL biosynthesis (Liu et al., 2016, 2019a; Byrne et al., 2017; Petersen et al., 2019), for GSL catabolism (Agerbirk and Olsen, 2015; Agerbirk et al., 2018), and as a gene donor for breeding and general biotechnology (Liu et al., 2017; Liu et al., 2019b).

In the present paper, we focus on the peculiar divergence of *B. vulgaris* in two genotypes of yet unsettled taxonomic status. These genotypes, named from a simple morphological character of first year rosette plants as the G-type (for glabrous) and the P-type (for pubescent) (Agerbirk et al., 2003), were reported to have contrasting geographic distribution but co-exist in a limited zone (Christensen et al., 2014; Heimes et al., 2016). The two types exhibit reduced interbreeding compatibility (Christensen et al., 2016) as well as contrasting chemistry (Byrne et al., 2017) and resistance properties (Badenes-Perez and Lopez-Perez, 2018; Christensen et al., 2019). The P-type with an Eastern geographical distribution is particularly rich in rare GSLs (Agerbirk et al., 2015). The entire genus seems to lack aliphatic GSLs but contains a surprising diversity of those derived from homoPhe and Trp (Olsen et al., 2016).

Several unsettled questions for the genus *Barbarea* and the tribe Cardamineae were raised by Olsen et al. (2016). These included the distribution in the tribe of a number of rare GSLs discovered in *A. thaliana* and *Barbarea*, GSL profiles in a wider range of *Barbarea* species, the profile of *Planodes virginica* (L.) Greene only investigated using classical methods, and the profile of the emerging model plant *C. hirsuta*. Other questions were related to conflicting literature, including the GSL profile of horseradish and a controversial historical report of aliphatic GSLs in *Barbarea*. Furthermore, recent molecular genetic studies had identified genes in *B. vulgaris* that could potentially be involved in biosynthesis of aliphatic (n-homoMet-derived) GSLs (Liu et al., 2016; Byrne et al., 2017; Wang et al., in press). Some of the anticipated GSLs were previously known to be organ-specific, making

further analysis of a range of organs relevant.

1.1. Comprehensive abbreviations for glucosinolates and precursor amino acids

Pioneering authors in Phytochemistry introduced a set of standard abbreviations for the structures identified in *A. thaliana* (Brown et al., 2003), and later extended the system ad hoc to additional structures in related species (Windsor et al., 2005). These Brown-Windsor abbreviations have gained almost universal popularity in papers dealing with *A. thaliana* GSLs and *Brassica* GSLs, which mainly overlap. An older paper used a somewhat different system (Daxenbichler et al., 1991). To extend to all known structures, the systems have been adapted here to cover all relevant structures, using a system as close to general organic chemistry nomenclature as possible but abbreviated and simplified. Following this tradition of structurally based abbreviations, we use the resulting comprehensive, systematic abbreviations consistently in the text. Most abbreviations should be intuitively recognizable by users of the Brown-Windsor abbreviations, e.g. IM for indol-3-ylmethylGSL, PE for phenethylGSL, 2mPr for 2-methylpropylGSL and 4mSOb for 4-methylsulfinylbutylGSL. A systematic explanation is provided in the figure legend (Supplementary Fig. S1).

For simplicity, we do not distinguish desulfoglucosinolates (dGSLs) from GSLs in the modified Brown-Windsor abbreviations, as the preceding "d" or similar would be confusing. For abbreviations written next to chromatographic peaks, which are all dGSLs in this work, the desulfo nature is implicit. In the text, we specify desulfo nature when relevant.

For amino acids, standard three-letter abbreviations are used, and for homologs we add the prefix 'homo', e.g. homoIle for homoisoleucine. For higher homologs, we use numbers rather than Greek prefixes for brevity, e.g. 2homoIle for dihomoisoleucine and 5homoMet for pentahomomethionine, and for indicating an unspecified higher homolog or a range of chain lengths, we replace the number with 'n' or the relevant numeral range, e.g. 1-2homoIle for mono-to dihomoisoleucine. Branched-chain amino acids (BCAAs) comprise Val, Leu, Ile and their homologs (Supplementary Fig. S1B). Glucose is abbreviated Glc.

1.2. Comprehensive numbers for glucosinolates

Another key condition for general comparison of GSL profile data is overview and consensus of scientifically known GSLs. Pioneers in cataloging botanical GSL diversity used a comprehensive numbering system, which has been used in four major Phytochemistry reviews (Fahey et al., 2001; Agerbirk and Olsen, 2012; Olsen et al., 2016; Blažević et al., 2020). This system assigns a fixed number to each scientifically accepted GSL (Blažević et al., 2020). We use this numbering system as a supplement in tables, figures and occasionally in text, in order to simplify comparison with the remaining literature. Occasionally, we use numbers only for some peaks for space considerations in figures and tables. Numbers in square brackets indicate less well-characterized structures not yet considered proven in any plant, i.e. those for which NMR characterization has never been reported, e.g. **[89]** for putative 9-(methylthio)nonyl GSL, 9mSn (Blažević et al., 2020). For dGSLs, a "d" is added to the numbers, such as d**11** for desulfo-**11**.

1.3. Scope and aims

The work presented here had three aims. The first aim was to strengthen the knowledge of GSL diversity in tribe Cardamineae by conclusive analysis of selected species and organs. For a comparative perspective, a distantly related outgroup (*Reseda*) was analyzed. This aim included some analytical optimization and extension of our library of relevant MS2 spectra. The second aim focused on the genus *Barbarea*, testing whether apparently absent aliphatic GSLs in *B. vulgaris* could be induced by various treatments. The third aim was to characterize accessions of *B. vulgaris* phylogenetically, to understand whether the investigated P-type and G-type material was representative for the species. The results are discussed case by case in relation to previous results on each species or genus, while a broader comparison and an evolutionary discussion are reported in a separate, accompanying review paper (Agerbirk et al., 2021).

2. Results and discussion

2.1. Library of mass spectra of standard desulfoglucosinolates

A central aim of this investigation was to confirm both the presence and absence of a wide variety of GSLs in individual samples. Like most previous authors in the field (e.g. Brown et al., 2003; Windsor et al., 2005), we converted GSLs to dGSLs by a well-established enzymatic on-column procedure. This conversion was done for two reasons. First, the availability of many authentic standards of dGSLs enabled peak identification supported by retention time (t_R) and MS2. Second, the increased specificity obtained by the desulfation procedure and the consequent clean nature of chromatograms simplified discovery of all GSLs in the sample and quantification based on UV-peak areas.

Ion trap MS2 spectra of all investigated dGSLs were recorded, as listed in Supplementary Fig. S2 with MS2 spectra of 68 dGSLs, and with MS data or reference to the literature for a further few. The general properties of the spectra, including use in isomer distinction, have been discussed previously (Agerbirk et al., 2014, 2015; Olsen et al., 2016; Pfalz et al., 2016; Pagnotta et al., 2017); observed fragment types are summarized in Table 1. Besides *m/z* value and t_R, dGSLs exhibited characteristic, fingerprint-like MS2 spectra, composed of a selection of the general fragment types that could be rationalized in terms of dGSL structure. The fragment codes (Table 1) are used in reported MS2 spectra in figures and the reported MS2 library.

In the rest of this section, only new developments are reported and

Table 1
Fragment types observed, alone or in combination, in ion trap MS/MS of desulfoglucosinolates (dGSLs) in positive mode with Na^+ or H^+ as cation.

Code[a]	Rationalization[b]
a	Cation-adduct of anhydroGlc, $C_6H_{10}O_5$ (Na^+ adduct at *m/z* 185) or an acyl derivative
b	Cation-adduct of thioGlc, $C_6H_{11}O_5SH$ (Na^+ adduct at *m/z* 219) or an acyl derivative
c	Loss of anhydroGlc (*m/z* 162) or an acyl derivative
d	Cation-adduct of Glc, $C_6H_{12}O_6$ (180) (Na^+ adduct at *m/z* 203)
e	Loss of aldehyde after aldol-like fragmentation of β-hydroxyls, forming desulfo methylGSL or isomer (Na^+ adduct at *m/z* 276) from usual dGSLs (e.g. d**24**, d**40**), or acylated counterpart from SGlc-acylated dGSLs (e.g. d**131**, d**132**).
f	Loss of water (*m/z* 18)
g	Loss of $H(SO)CH_3$ (*m/z* 64)
h	Loss of thioGlc (*m/z* 196) or an acyl derivative
i	Loss of Glc (*m/z* 180)
j	Cation-adduct of gluconolactone, $C_6H_{10}O_6$ (Na^+ adduct at *m/z* 201)
k	Loss of gluconolactone (*m/z* 178)
m	Loss of mercaptomethane, CH_3SH (*m/z* 48)
n	Loss of oxo-thioGlc (*m/z* 212) from methylsulfinylalkyl dGSLs etc. (e.g. d**73**, d**64**, d**65**, d**5**)
Misc.	Various other fragments, including HO and OCH_3 radicals from homolytic cleavage of bond to *N* (in d**47**, d**138**, d**154**), occasional second loss of H_2O (e.g. d**131**), occasional loss of NH_2OH (*m/z* 33) from $[M+H]^+$ (e.g. d**68**), occasional loss of H_2S (*m/z* 34) from dGSLs with a side chain hydroxy group (d**42**, d**140**, d**141**), anhydro-rhamnose from rhamnosylated dGSLs (d**109**, d**110**), and resonance-stabilized arylmethylene type fragments from indole dGSLs (e.g. d**43**, d**48**), from some substituted benzylic dGSLs (e.g. d**114**, d**152**) and from 6′-isoferuyloylated dGSLs (d**131**, d**132**).

[a] Fragment types have codes as in the original reports, in order of discovery/characterization. The letter "l" was avoided. Types k, m and n are reported in this work. Types d, i, j and k require rearrangement (Olsen et al., 2016).

[b] For simplicity, cation adducts of the remaining part of M after neutral loss of a smaller molecule X is simply termed "Loss of X". Acyl derivatives are only specified if observed in at least one case. Glc, glucose.

discussed, with focus on distinction of isomers by MS2. The added authentic standards since the 2016 report (Olsen et al., 2016) supplied almost all missing relevant standards. Two standard GSLs, 2h2mPr (**31**) and 2h2mBu (**29*S***), were unfortunately not obtained for the ion trap MS experiments even though they have been reported from the tribe Cardamineae, but were of types that we expected to be detectable. In addition to the previously characterized seeds of *Arabis soyeri* Reut & Huet ssp. *subcoriaceae* (Gren.) Breistr., *Sinapis alba* ssp. *mairei* (Lindb.) Maire, *Sinapis arvensis* L. and various *Barbarea* types and species, which are publicly available at Nordgen (https://nordic-baltic-genebanks.org), we describe two other reference seeds, *P. virginica* and *N. officinale*, available at Nordgen and commercially, respectively (Section 4.2.). These reference seeds contain a range of long-chain methylsulfinylalkylGSLs and methylthioalkylGSLs.

In the former paper (Olsen et al., 2016), t_R and MS of 70 proven or suggested dGSLs plus some candidate analytes were recorded, but eight of them referred to GSLs that have never been fully characterized by NMR in any scientific paper: 9mSn (**[89]**), suggested "2-hydroxy-8-(methylsulfinyl)octylGSL" ("**150**"), 1hIM (**[154]**), **x1**, **x2** and **x4** with uncertain substitution pattern, and "**c2**" and "**c3**" with uncertain substituents. Hence, only 62 fully proven GSLs were covered. The total number of conclusively known GSLs listed in a recent review was 88 (Blažević et al., 2020), since then one additional conclusive report has appeared (Montaut et al., 2020). In an attempt to approach the entire number (89) of universally proven GSLs, with focus on those that were relevant for the tribe, we acquired some additional references. Due to occasional literature suggestions of two unsaturated GSLs in the tribe, 4-(methylthio)but-3-enylGSL (4mSbuen) and the corresponding sulfinyl 4mSObuen, they were relevant for validation of our conditions. Furthermore, we included 3-(methylsulfinyl)propylGSL (3mSOp) and seeds of the Col-0 genotype of *A. thaliana* as a reliable reference material for medium chain length Met-derived GSLs, and for hydroxyalkyl GSLs and various benzoyl esters (Reichelt et al., 2002; Brown et al., 2003). Although *A. thaliana* Col-0 seeds are rather well investigated, the presence of 5BzOp (**117**) had to be interpreted from the context (Reichelt et al., 2002), as its presence in the Col-0 ecotype was not explicitly reported in the pioneering papers (Reichelt et al., 2002; Kliebenstein et al., 2001). The MS2 spectrum of apparent d**117** is hereby reported from the Col-0 ecotype (Supplementary Fig. S2). Finally, we included seeds of *Reseda odorata* L. and *Iberis amara* L. for specific well-established GSLs. Due to column change, instrument repair and a slight adjustment of the gradient between this and the previous report, it was not surprising that t_R values had generally changed a bit, in this case towards earlier values (Table 2).

From the newly acquired references, we concluded that the well-established *A. thaliana* GSLs were detectable at our conditions and that double bonds had little influence on t_R of dGSLs: 0.2 min delay for desulfated 4mSbuen (**83**) vs. 4mSb (**84**) and coelution at our conditions of desulfated 4mSObuen (**63**) and 4mSOb (**64**) (Table 2). Two recently discovered compounds (d**151**, d**152**), corresponding to polysubstituted BZ (Pagnotta et al., 2017) and a related dGSL (d**15**) were not properly defined in the previous list and two were accidently interchanged (d**15** vs. d**152**) but are correctly shown here. The MS data and t_R of each new standard is listed along with data for a broad selection of previously studied references. They all showed expectable relation between structure and t_R. Some authentic standards from a previous paper (Olsen et al., 2016) were not included in the present list because the operation of the ion trap MS instrument abruptly had to be discontinued for financial reasons, but MS data from both lists are included in Supplementary Fig. S1. Due to the discontinuation of ion trap HPLC-MS, a UHPLC-QToF MS/MS method was developed (Section 4.4.) and applied for some final experiments, confined to Supplementary Fig. S4 and a distinct part of Supplementary Fig. S5, meaning that results from the UHPLC-QToF MS/MS method are not mixed with results of the ion trap-based method. The UHPLC-QToF MS/MS method will be further described and discussed separately in a future publication (Agerbirk and

Table 2
Retention time in HPLC (t_R) and MS information of selected desulfoglucosinolates (dGSLs), along with the evidence for the identity of the reference compound or reference material. Those with an asterisk (*) after the number were only tentatively identified.

Global number and acronym		Masses of dGSLs (Da)[a]		t_R of dGSLs (min)	Evidence for dGSL identity[b]
GSL		Desulfo $[M+H]^+$	Desulfo $[M+Na]^+$		
107	Pren	280	302	1.5	NMR
56	1mEt	282	304	2.4	NMR
12	Buen	294	316	2.8	NMR
62	2mPr	296	318	3.6	NMR
61	1mPr	296	318	4.5	NMR
[42]	3hPr	298	320	1.1	Col-0[c]
57*R*	2hmEt	298	320	1.4	NMR
101	Peen	308	330	5.3	NMR
24*R*	R2hBuen	310	332	1.3	NMR
24*S* *	S2hBuen	310	332	1.4	*Crambe*[d]
[26]	4hBu	312	334	1.2	Col-0[c]
30	1hmPr	312	334	2.4	NMR
31	2h2mPr	312	334	–	–
58	3mPe	324	346	7.2	NMR
95*	3mSp	328	350	4.2	*Iberis* [d], hrMS, MS2
11	BZ	330	352	5.4	NMR
141	3hmPe	340	362	4.0	NMR
149	2h3mPe	340	362	5.3	NMR
83	4mSbuen	340	362	5.8	NMR/Vis92
84	4mSb	342	364	5.6	NMR
73	3mSOp	344	366	1.2	Comm., MS2
105	PE	344	366	6.6	NMR
23	4hBZ	346	368	2.6	NMT
22	3hBZ	346	368	3.1	NMR/Pag17
63	4mSObuen	356	378	1.4	NMR/Ior08
64	4mSOb	358	380	1.4	NMR
40*R*	EBAR	360	382	4.4	NMR, *B. vulg.*P[c]
40*S*	BAR	360	382	4.9	NMR, *B. vulg.*G[c]
140	4hPE	360	382	4.8	NMR
148	3hPE	360	382	5.4	NMR
43	IM	369	391	6.0	NMR
72	5mSOp	372	394	2.2	Col-0[c]
139*R*	4hEBAR	376	398	2.1	NMR
139*S*	4hBAR	376	398	2.2	NMR
142*R*	3hEBAR	376	398	2.5	NMR
151	4h3moBZ	376	398	3.0	NMR/Pag17
87	7mSh	384	406	8.6	Col-0[c]
[154]	1hIM	385	407	6.4	MS2/Pfa16
67	6mSOh	386	408	4.7	Col-0[c]
15	3,4moBZ	390	412	5.8	NMR/Pag17
92	8mSo	398	420	9.5	NMR
48	4moIM	399	421	6.7	NMR
47	1moIM	399	421	7.6	NMR/Bla20
66	7mSOh	400	422	5.6	NMR
[10]	3BzOp	402	424	7.6	Col-0[c]
x1	3/4hmoEBAR	406	428	2.4	NMR
152	4h3,5moBZ	406	428	4.7	NMR/Pag17
69	8mSOo	414	436	6.5	NMR
5	4BzOb	416	438	8.3	Col-0[c]
114	3,4,5moBZ	420	442	6.2	NMR/Pag17
68*	9mSOn	428	450	7.3	MS2 (Fig. 1)
138	1,4moIM	429	451	8.0	NMR
80	8mSO$_2$o	430	452	6.8	NMR
117	5BzOp	430	452	8.9	Col-0[c]
79	9mSO$_2$n	444	466	7.7	NMR
127	6′Bz 4mSb	446	468	9.8	Col-0[c]
77	10mSO$_2$d	458	480	8.4	NMR
109*	2RhaOBZ	492	514	5.9	*R. odorata* [d]
129	6'iF PE	520	542	9.8	NMR, *B. vulg.*P[c]
125	6′Bz 4BzOb	520	542	10.9	Col-0[c]
131*R*	6'iF EBAR	536	558	9.0	NMR, *B. vulg.*P[c]

(*continued on next page*)

Table 2 (*continued*)

Global number and acronym		Masses of dGSLs (Da)[a]		t_R of dGSLs (min)	Evidence for dGSL identity[b]
GSL		Desulfo $[M+H]^+$	Desulfo $[M+Na]^+$		
131*S*	6'iF BAR	536	558	9.1	NMR, *B. vulg.*G[c]
130	6'iF IM	545	567	9.5	NMR, *B. vulg.*P[c]
132*R*	6'iF 4hEBAR	552	574	7.9	NMR, *B. vulg.*P[b]
134*	$(Sbu)_2$	653	675	6.9	*Eruca*[d], *m/z*

[a] Many dGSLs are isobaric, as listed here or in Olsen et al. (2016) (in case of 346, 407 and 412 for $[M+Na]^+$).

[b] hr, high resolution, Comm., commercially obtained reference, r.n.s., results not shown. References for spectroscopic identification are as reported by Olsen et al. (2016) or in notes below, except for the following: Vis92, Visentin et al. (1992); Ior08, Iori et al. (2008); Pfa16, Pfalz et al. (2016); Pfa17, Pfalz et al. (2017); Bla20, Blažević et al. (2020).

[c] Col-0 indicated as evidence means that seeds of the Col-0 accession of *Arabidopsis thaliana* was used as reference material based on a thorough published analysis (Brown et al., 2003), *B. vulg.* means that publicly available reference seed accessions of *Barbarea vulgaris* (P-type or G-type as indicated) were routinely used as reference material based on a published analysis (Agerbirk and Olsen, 2011).

[d] Putative **24*S*** was a major peak in *Crambe maritima* L. foliage, putative **95** (confirmed with high resolution MS and MS2 using QToF MS) was a minor peak in seeds of *Iberis amara*, together with the major **73** confirmed by comparison with the authentic reference, putative **109** was the dominant peak in seeds of *Reseda odorata* (Fig. 9H), putative **134** was a minor peak in foliage of *Eruca* sp. from a food store.

Pattison, unpublished). Unless specifically indicated to be QToF MS, all discussion of MS fragmentation in this paper refers to ion trap MS.

The MS fragmentation of corresponding dGSLs with or without a double bond at the side chain thiofunctionality was complex (Fig. 1). It was well-established that ω-(methylsulfinyl)alkylGSLs showed a neutral loss of 64 Da (H_3CSOH) from the Na^+ adduct, and mainly formed MS2 Na^+ adducts with fragments containing the polar sulfinyl group. For example, the MS2 spectum of the Na^+ adduct of desulfo 4mSOb showed major fragments corresponding to loss of H_3CSOH (type g) and loss of anhydroGlc (type c) (Fig. 1A). However, the analog desulfo 4mSObuen with a double bond near the sulfinyl group did not exhibit the loss of H_3CSOH but was entirely dominated by loss of anhydroGlc (type c) in MS2 (Fig. 1B). Hence, the extra double bond in desulfo 4mSObuen inhibited fragmentation next to the methylsulfinyl functionality.

The situation was opposite for the analogous pair of methylthio functionalised dGSLs. In this case, the saturated desulfo 4mSb solely fragmented at either side of the thioGlc sulfur, and Na^+ adducts mainly formed with the anhydroGlc and thioGlc fragments, as expected due to the low polarity of the side chain functionality (Fig. 1C). However, the unsaturated desulfo 4mSbuen exhibited a fragment (m/z = 152) that seemed to be due to loss of H_3CSH after the loss of anhydroGlc (type c+m) (Fig. 1D). In order to test this hypothesis, we checked the fragmentation of the $[M+H]^+$ adducts, which were also formed for both these dGSLs. In both cases, we observed the combined loss of anhydroGlc and H_3CSH (Fig. 1 E+F). From unsaturated 4mSbuen, we even observed loss of H_3CSH (type m) from the pseudomolecular ion $[M+H]^+$ (Fig. 1F). The MS2 spectrum of the Na^+ adduct of this dGSL, d**83**, was also reported by Kakizaki et al. (2017) based on fragmentation in orbitrap MS. However, in that spectrum only the major fragments at m/z 185 (type a) and 219 (type b) were detected, suggesting that the alternative fragmentation (Fig. 1D) depends on the method or conditions of fragmentation.

One more fragment type was discovered by systematic comparison of MS2 spectra: a loss of 212 that was always associated with the known

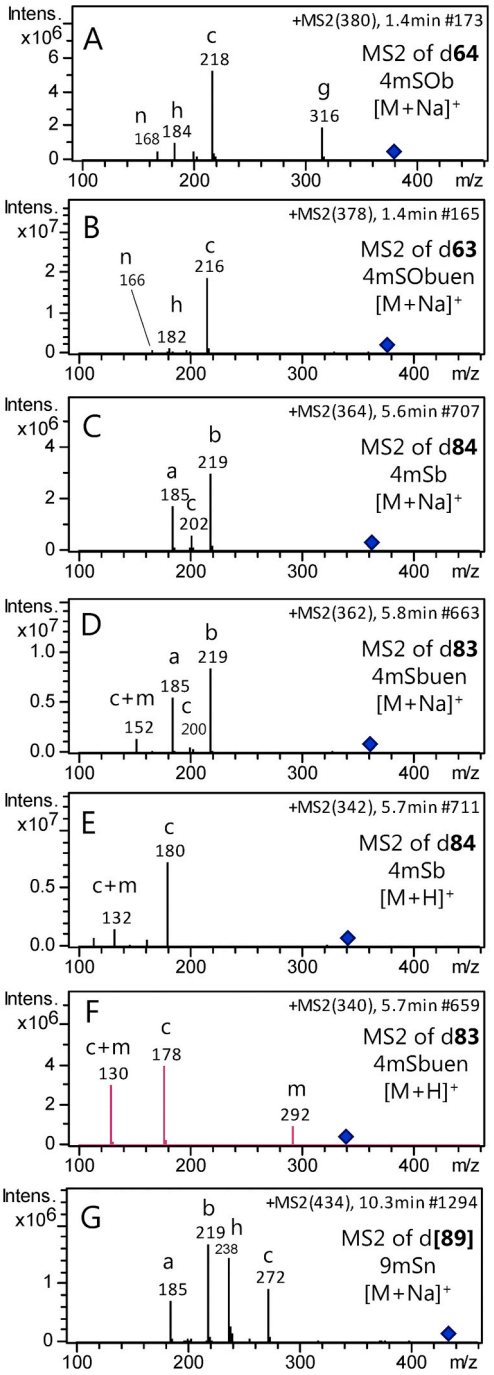


Fig. 1. MS2 spectra of pairs of desulfoglucosinolates with and without a side chain double bond. Four short chain desulfoglucosinolates were investigated, including Na^+ adducts of all (A–D) and in addition H^+ adducts of the methylthio substituted (E–F), as indicated in each spectrum. The desulfo derivative of the putative 9mSn (**[89]**), poorly characterized in the literature, was also investigated (G).

loss of 196 (type h, loss of thioGlc) typical of dGSLs with polar side chains. We noticed that the mass 212 was 16 Da higher than the mass of thioGlc and mainly occurred for dGSLs with a methylsulfinylalkyl side chain. Therefore, we speculated that the loss of 212 could be a loss of oxidized thioGlc, probably due to oxidation of the mercapto group, formed by reaction with the sulfinyl group. As the fragmentation was characteristic of an entire homologous series, it was given a letter code (n) (Table 1). This fragment type was also observed for some methylsulfonylalkyls (e.g. d**64** and d**80**) and for unsaturated d**63** (Fig. 1A–B).

Hence, this calibration work had identified two new diagnostic fragmentations in the side chain of specific dGSLs, loss of oxothioGlc from methylsulfinylalkyls in general and loss of H_3CSH from some methylthioalkyls and –alkenyls in MS2. The loss of H_3CSH was not observed in corresponding MS2 spectra of long chain methylthioalkyl dGSLs such as 7mSh and 8mSo (d**87**, d**92**), so this specific fragmentation was not general for this structural type. However, it provided an extra means of distinguishing the three isobaric dGSLs from 4mSbuen, 3hmPe and 2h3mPe (d**83**, d**141** and d**149**), all at $m/z = 362$ for $[M+Na]^+$ and all suggested from the tribe, although the suggestion of **83** in tribe Cardamineae is not sufficiently evidenced (Agerbirk et al., 2021). An MS2 spectrum of the very long chain desulfo 9mSn (d**[89]**) was also recorded (Fig. 1G), from previously reported analysis of seeds of *Rorippa amphibia* (L.) Besser (Olsen et al., 2016), as a contribution to the documentation for this putative natural GSL. The fragmentation of d**[89]** fully agreed with the proposed structure, as it was identical to the fragmentation of the lower homolog d**92**, with the actual masses of the side-chain containing fragments, h and c, shifted 14 Da towards higher m/z values.

Because we searched for GSLs that are acylated on the thioGlc moeity, it was of particular interest to test whether the known thioGlc acylated GSLs from *A. thaliana* seeds (Reichelt et al., 2002) were detectable at our conditions (Fig. 2). This was the case for dGSLs from 6′Bz 4mSb (**127**) and 6′Bz 4BzOb (**125**), and the MS2 from Na^+ adducts of both showed a dominant fragment at m/z 323, interpreted as the Na^+ adduct of benzoyl thioGlc. Similarly, it was known that 6′-isoferuloyl derivatives of dGSLs (Supplementary Fig. S1A) form a dominant MS2 fragment corresponding to an isoferuloyl thioGlc Na^+ adduct at m/z 395 (Agerbirk and Olsen, 2011). Both kinds of dGSLs showed affinity for the DEAE Sephadex column material even after analytical desulfation and exhibited delayed elution, but complete elution was obtained with our protocol that includes thorough elution (five times one mL) after desulfation (Agerbirk and Olsen, 2011). Hence, thioGlc acylated glucosinolates in general could be detected and identified by MS2 at our conditions, allowing search in seeds of other species. We noticed the isomeric nature of a benzoylated GSL (**125**) from *A. thaliana* and an isoferuloylated GSL (**129**) from the genus *Barbarea* (Fig. 2), underlining the relevance of authentic standards and efficient chromatographic separation even for quite complex GSL structures.

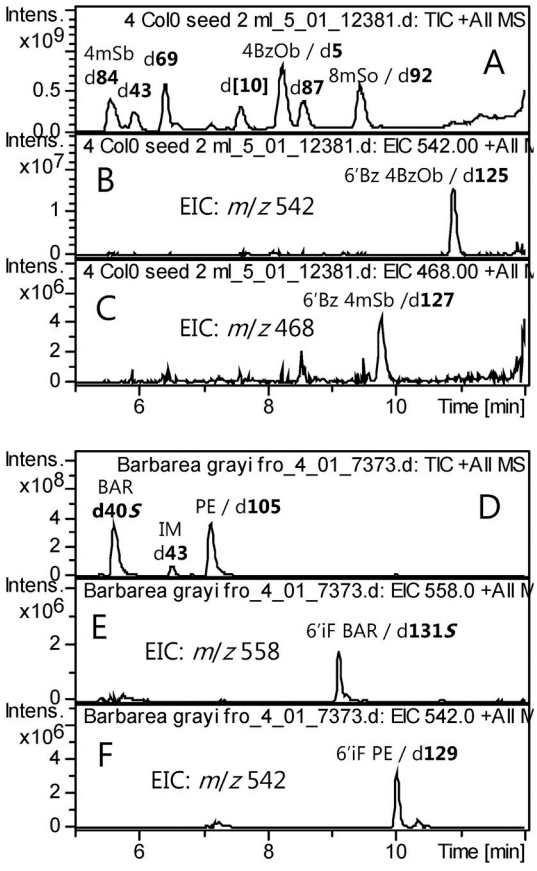


Fig. 2. Detection of thioglucose-acylated glucosinolates (GSLs) by HPLC-MS of desulfated derivatives prepared from the indicated species. (A–C) Analysis of seeds of *Arabidopsis thaliana* Col-0 used as reference material for characteristic GSLs. Shown are the total ion chromatogram (A) and extracted ion chromatograms for sodium adducts of desulfo 6′Bz 4BzOb (d125) (B) and desulfo 6′Bz 4mSb (d127) (C). (D–F) Analysis of seeds of *Barbarea grayi* for dominating GSLs. Shown are total ion chromatograms (D), and extracted ion chromatograms for sodium adducts of desulfo 6'iF BAR (d131*S*) (E) and desulfo 6'iF PE (d129) (F).

All in all, this standardization work added 11 new standards to our library, including three from *A. thaliana* at present only partially characterized (**[10]**, **[26]**, **[42]**). From a combination of data in Olsen et al. (2016) and here, our analysis protocol allowed us to conclude the presence or absence (below the limit of detection) of at least 62 + 8 = 70 fully characterized GSLs (according to Blažević et al., 2020), plus five partially characterized but generally accepted **[10]**, **[26]**, **[42]**, **[89]** and **[154]** that are intermediates in well characterized biosyntheses or part of a well characterized homologous series. In addition, our conditions allowed detection of some not yet fully identified GSL-like compounds so far only known as trace peaks from *A. thaliana* mutants ("*N*-Hex-I3M" and "*N*-X-I3M"), *B. vulgaris* ("**x1**", **x2**″, "**x4**") and *R. amphibia* ("**150**") as specified by Pfalz et al. (2016) and Olsen et al. (2016). Besides m/z value and t_R, MS2 spectra were informative and allowed recognition of, e.g., methylsulfonyls, methylsulfinyls, short chain methylthios, β-hydroxyls and acylation at the thioglucoside. A known and a tentatively known GSL with a negative charge in the side-chain, **112** and **[155]** (Supplementary Fig. 1C) would not be included since they are incompatible with elution with pure water after desulfation, and acetyl groups can be lost during desulfation (Blažević et al., 2020). All other known or suggested GSLs would be expected to be included in our analysis, if present.

2.2. Glucosinolate profiles in the tribe Cardamineae

The selected tribe Cardamineae species were investigated for GSLs with the overall evolutionary history in mind. In particular, we searched for similarities to the deviating genus *Barbarea* and *C. pratensis*. The characteristic GSLs in *Barbarea* include the homoPhe-derived "glucobarbarin" (BAR, **40*S***), which is (*S*)-2-hydroxy-2-phenylethylGSL, epiglucobarbarin (EBAR, **40*R***), which is the epimer (*R*)-2-hydroxy-2-phenylethylGSL, the Trp-derived IM (**43**), and a number of derivatives. Among the derivatives are the disubstituted 1,4moIM (**138**) (usually only detectable in roots), phenolic derivatives of both PE, BAR and EBAR, and isoferuloyl derivatives of the same and of IM (only known from seeds) (Supplementary Figure 1). Interestingly, the "innovative" *C. pratensis* also accumulates 1,4moIM (Olsen et al., 2016). So far unique GSLs in *C. pratensis* are two hydroxylated GSLs with a 3-methylpentyl skeleton, 2h3mPe (**149**) and 3hmPe (**141**), both predicted to be biosynthesized from 2homoIle (Agerbirk et al., 2010; Olsen et al., 2016). One of them (**149**) was likewise confined to roots (Agerbirk et al., 2010). These recently discovered GSLs were found in some American and commercial accessions, while Danish accessions of *C. pratensis* contained a variety of other BCAA-derived GSLs and Phe or Tyr-derived GSLs (Agerbirk et al., 2010).

A variety of organs were investigated whenever possible, in order to maximize the chance of detecting critical structures. In tables in this section, we consistently distinguished tentative from conclusive identifications based on an MS2 dependent criterion: For each species, conclusive identification of each reported GSLs was concluded when two independent characteristics of a peak matched an authentic standard (Blažević et al., 2020): (1) correct t_R of the relevant m/z value was

observed in extracted ion chromatograms, and (2) the correct MS2 spectrum was observed. Correctness of t_R and MS2 were based on comparison with an authentic standard (Table 2). When conclusive identification was obtained from one organ, minor peaks from other organs were accepted as conclusive as well. In contrast, tentative identification (Blažević et al., 2020) was concluded when a clearly distinguishable peak was observed but the MS2 criterion was not fulfilled for any organ of the species, or when reasonable t_R and MS2 was observed but an authentic standard was not available. In either case, the tentative nature was indicated by labelling the data with an asterisk (*). It follows from these criteria that identification of GSLs for which authentic standards were not available were by definition tentative, even for members of a homologous series with expectable MS2. In those cases, the GSL number in the entire table was labelled with an asterisks, essentially 3mSp (**95***), 9mSOn (**68***), 5mSp (**94***) and 9mSn (**[89]** *), S2hBuen (**24*S***) and 2RhaOBZ (**109***). For individual levels of GSLs of this small group, an additional asterisk at individual data points means that also MS2 confirmation was lacking, while individual data points without an asterisk represent data validated by expectable MS2 for the side chain type.

Documentation for botanical identity and specification of origin followed recent recommendations (Zidorn, 2017) as far as possible (Section 4.2.).

2.2.1. Endemic Australian Barbarea species

Our previous investigations of *Barbarea* GSL profiles had focused on species from the Northern temperate zone (Agerbirk et al., 2003a, 2015), while many other species of the genus appeared to be poorly known with respect to chemistry. For the purpose of investigating distantly related species, we turned to two endemic Australian species, *Barbarea australis* Hook f. and *Barbarea grayi* Hewson. Both species contained expectable homoPhe and Trp derived GSLs in leaves and roots (Table 3), and no surprising GSLs were observed. Rather, PE and the usual derivatives BAR and EBAR were dominating GSLs, along with expectable indole GSLs.

Seed GSLs of *B. australis* had already been reported (Agerbirk and Olsen, 2011), and included most of the GSLs from leaves and roots (except substituted Trp-derived) as well as low levels of 6′-isoferuloyl derivatives of the major seed GSLs PE and BAR. Seed GSLs of *B. grayi* were analyzed for this report and contained the expected isoferuloylated GSLs known from the genus, but at rather low levels (Table 3). A large number of putative aliphatic GSLs and derivatives of the major GSLs were searched for but not found (Table 3).

2.2.2. The genus Planodes

The genus *Planodes* comprises two species only. Seeds of the small weedy plant *P. virginica* had been investigated previously using now historical methods (Gmelin et al., 1970). Those authors reported the presence of EBAR (**40*R***). 7mSOh (**66**) and 8mSOo (**69**), two of which were deduced for the first time in that paper, and one or more lipophilic GSLs that were not identified. We wished to repeat and expand this experiment, testing all major plant parts, with focus on searching for similarities with and contrast to *Barbarea* and *Cardamine* species. In particular, we wanted to search for oxidized derivatives of homoPhe derived GSLs, for distubstituted Trp-derived in roots, for acyl derivatives in seeds, and for GSLs derived from BCAAs, as many of these kinds of structures were potentially due to *de novo* evolution in one or more species in the tribe. Likewise, we wished to know the full complement of n-homoMet-derived GSLs in the species, as this profile was a candidate for an ancestral profile of the *Barbarea* species.

The historical report on the dominant glucosinolates in *P. virginica* seeds was fully confirmed (Fig. 3A). Of particular interest was the dominance of EBAR that also occurs in most *Barbarea* species (Table 4). In addition, minor levels of BAR and NAS were discovered. Of Trp-derived GSLs, only the parent structure IM was detected, in contrast to the prominence of monosubstituted indoles in roots of most other known crucifers. Clearly, this species did not accumulate the peculiar disubstituted indole, 1,4moIM, known from roots of one *Barbarea* species.

In addition to the previously known dominating methylsulfinyl alkyls with chain length 7–8, a wide range of chain lengths were found (Table 5), supported by coelution with authentic standards except in the case of 9mSOn (**68**). For 5mSOp (**72**), the only authentic standard available was *A. thaliana* seeds, which was considered reliable. The retention times of the homologs increased in an expectable way for a homologous series (Supplementary Fig. S3C). The characteristic MS2 fragmentation (including loss of CH_3SOH) was observed for all major peaks including the tentatively identified 9mSOn (Supplementary Fig. S3D). From the latter fragmentation, isomeric hydroxylated methylthioalkyls could be ruled out. The 'bell-shaped' level distribution of chain lengths (Table 5) was similar to distributions in some other species, as expected from the properties of MAM-enzymes that are responsible for the chain elongation (Textor et al., 2007; Kumar et al., 2019; Petersen et al., 2019).

The more lipophilic methylthioalkyl glucosinolates were mainly detected in seeds (Table 5). Only two were identified with certainty, 8mSo from comparison with an authentic standard, and 7mSh from comparison with *A. thaliana* seeds, considered a reliable reference material. However, additional members of the entire homologous series from n = 5 to n = 9 were tentatively detected, supported by reasonable t_R compared to that of 8mSo, 7mSh and of 4mSb available as standards. Due to the wide range of GSLs confirmed in *P. virginica* seeds and their public availability at Nordgen, we imagined that they could be useful as

Table 3
Glucosinolate profiles of *Barbarea australis* and *B. grayi*.

	N	Total	SD	PE	BAR	EBAR	IM	4hIM	4moIM	1moIM	6'iF PE	6'iF BAR
				105	**40*S***	**40*R***	**43**	**28**	**48**	**47**	**129**	**131*S***
Species (accession) Organ							µmol/g dry wt.					
B. australis[a]												
Roots[b]	2	6.9	3.0	3.7	2.3	0.02	0.51	0.01*	0.33	n.d.	n.d.	n.d.
Leaves	2	43	9	2.3	39	0.45	1.4	n.d.	0.07	n.d.	n.d.	n.d.
B. grayi												
Roots[c, b]	3	38	11	17	13	0.17	6.0	0.01*	1.4	0.04*	n.d.	n.d.
Leaves[c]	3	75	20	15	56	0.63	3.2	n.d.	0.25	n.d.	n.d.	n.d.
Seeds[c, d]	1	92	n/a	46	44	0.41	2.4	0.04*	0.12	n.d.	0.18	0.10

N is number of replicates, SD is standard deviation, other data are mean levels.

[a] Results for accessions B46 and B47 were comparable and are pooled.

[b] Minimum estimates for roots due to difficulties in harvesting clean, undamaged roots.

[c] Searched for by HPLC-MS but not found: BZ, 4hBAR, 4hEBAR, 3hEBAR, 4hPE, 3hPE, 3/4mhPE (**x2**), 3/4mhEBAR (**x1**), 1,4moIM, methylsulfinylalkyls (n = 3–10), methylthioalkyls (n = 7–9), Buen, Peen, 1mEt, 1mPr, 2mPr, 3mPe. GSL data with an asterisk indicate lack of MS2 confirmation of analyte identity due to low level.

[d] In addition, traces of 6'iF EBAR (**131*R***) and 6'iF IM (**130***) were detected.

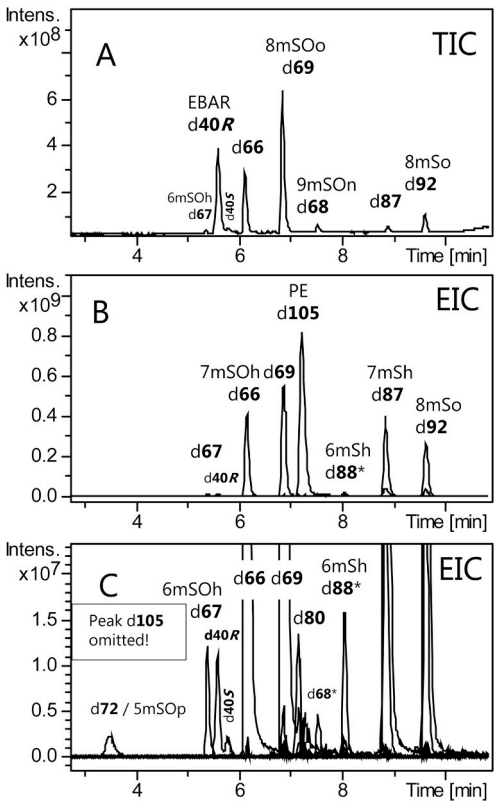


Fig. 3. HPLC-MS chromatograms of desulfoglucosinolates (dGSLs) prepared from glucosinolates (GSLs) in *Planodes virginica* (A) and *Nasturtium officinale* (B–C) seeds, showing qualitative similarities and quantitative contrasts. Major peaks (B) from *N. officinale* revealed many of the same GSLs as in A, but levels of EBAR (**40*R***) were much lower while levels were much higher for the biosynthetic precursor PE (**105**). A focus on trace peaks from *N. officinale* (C) revealed sharp peaks representing a range of minor constituents. Due to the closely eluting peaks, the latter chromatograms (B–C) were made by combining extracted ion chromatograms corresponding to $[M+Na]^+$ of the indicated dGSLs. In C, the *m/z* 366 signal of d**105** was omitted to allow visualization of minor coeluting peaks. An asterisk after a peak number indicates tentative identification. HPLC-MS conditions as in Olsen et al. (2016). TIC, total ion chromatogram, EIC, extracted ion chromatogram.

reference material and compiled various illustrative chromatograms (Supplementary Fig. S3). Pioneers in the study of *P. virginica* also confirmed the *R*-configuration of the chiral sulfur in the side-chain methylsulfinyl groups (Gmelin et al., 1970), which adds further reliability to this reference material.

Since *P. virginica* shared high levels of EBAR with *Barbarea*, some other characteristic *Barbarea* GSLs were searched for (Table 4). ThioGlc-acylated GSLs were searched for but not found. Similarly, hydroxyl and methoxyl derivatives of the major EBAR were searched for. A possible trace of the 4-hydroxy derivative 4hEBAR (**139*R***) was detected in many samples at very low levels but not confirmed by MS2 using ion trap HPLC-MS. As 4-hydroxylation is seasonally regulated in *Barbarea* (Agerbirk and Olsen, 2015), a new set of samples were collected earlier in the next year (accession 2), in an attempt to confirm or reject the presence of 4hEBAR. The putative trace of 4hEBAR was also observed in this second accession, but of the same very low level as in the first accession and still MS2 confirmation was not obtained by ion trap HPLC-MS. If 4hBAR or 4hEBAR were present, it would be the first occurrence outside the genera *Barbarea* and *Arabis*. In order to settle the uncertainty, we carried out more sensitive UHPLC-QToF MS/MS with higher mass- and chromatographic resolution. At these conditions, authentic desulfo 4hBAR and 4hEBAR mainly resulted in proton adducts after loss of water; $[M\text{-}H_2O+H]^+$. Re-analysis of three *P. virginica* leaf samples of accession 2 and seeds of accession 1 demonstrated that 4hEBAR was present at very low levels, supported by matching t_R and MS2. The high-resolution mass matched the expected (found: 358.0953 (mean, N = 4), calculated for $C_{15}H_{20}O_7NS^+$ ($[M\text{-}H_2O+H]^+$): 358.0955). Critical MS2 features supporting the identification was loss of anhydroGlc and a characteristic major loss of *m/z* 220, plus a combined loss of anhydroGlc and hydroxylamine (Supplementary Fig. S4A-C). In the seed sample, which had the highest level, loss of hydroxylamine from the $[M\text{–}H_2O\text{–ahGlc}+H]^+$ ion was also confirmed (Supplementary Fig. S4E). We concluded that 4hEBAR had been detected conclusively in two accessions of *P. virginica*, but at rather low levels (Table 4). We did not detect any other known derivatives of EBAR (the 3-hydroxy derivative 3hEBAR (**142*R***), the 4-methoxy derivative 4moEBAR (**50*R***) or the 3/4 methoxy-hydroxy derivative "**x1**") in any sample. As *P. virginica* shared chain elongated Met derived GSLs with *A. thaliana*, some rare and characteristic GSLs from this species were searched for but not found (hydroxyalkyls and their benzoyl esters). This negative result was likewise obtained for two more species discussed below (Table 5).

We also searched for GSLs derived from BCAAs, because their distribution in the tribe was poorly known. No GSL derived from non-chain elongated BCAAs was detected (Table 5). However, a peak corresponding to an isomeric hydroxyhexylGSL was detectable at minute levels, with a further isomer at close to trace levels (Supplementary Figure 4H, panels H1–H3). The sum formula was supported by QToF-MS (found 340.1424, calculated for $[C_{13}H_{26}O_7NS]^+$: 340.1425) (panel H1). Two such isomers are known, **141** and **149**. The t_R of the unknown dGSLs were far from that of d**141** but the major was close to d**149**. However, a reproducible slight (0.1 min) mismatch with d**149** was observed using UHPLC, confirmed by spiking that resulted in a split-peak (results not shown). A quite characteristic MS2 ion trap spectrum, with six major fragment ions, was confirmed in several samples (Supplementary Fig. S3G). The ion trap MS2 included an e-fragment (Table 1), which is indicative of a β-hydroxylated dGSL (Olsen et al., 2016). We conclude the presence of a hitherto unknown isomeric β-hydroxyhexyl GSL at levels around 0.01 μmol/g dry wt. in all tested plant parts of both *P. virginica* accessions. Possible identities could be a straight chain isomer or "2h4mPe" derived from 2-homoLeu, not yet known to science. Several structurally related known GSLs were searched for but not found, including the parent GSL 3mPe, the δ-hydroxylated isomer 3hmPe and lower homologs (Table 5).

2.2.3. The genus Rorippa

The genus had been investigated in detail earlier (Olsen et al., 2016), except seeds of *R. sylvestris*. Here, the seeds were investigated with the aim to search for seed-specific acylated GSL (Table 5). Green parts and flowers of the sampled individual matched the previously reported profiles (results not shown). The GSL profile of the seeds was dominated by 6homoMet-derived 8mSOo (**69**) and the corresponding methylthio precursor 8mSo (**92**), i.e. the "C8" chain length group corresponding to six rounds of chain elongation in the early biosynthesis, in agreement with the previously reported leaf and root profiles. From visual comparison of chain-length distributions in the common table (Table 5), the distribution in *R. sylvestris* was obviously more narrow than seen for *Nasturtium* and *Planodes*, suggesting a more stringent, processive chain elongation system (Textor et al., 2007; Kumar et al., 2019; Petersen et al., 2019). Acylated GSLs, BCAA-derived, Trp-derived or Phe-derived GSLs were not detected (Table 4, Table 5) but Trp-derived GSLs are confirmed in other organs of the same species (Olsen et al., 2016).

2.2.4. Watercress

The tribe Cardamineae includes two horticultural crops, one of which is watercress (*N. officinale*). In a previous paper, two contrasting GSL profiles of watercress were reported (Agerbirk et al., 2014). However, a search for acyl derivatives in seeds was lacking. We extended the seed analysis with two further accessions and confirmed the previously reported seed GSL profile for the species (Fig. 3B–C, Supplementary Fig.

Table 4
Total glucosinolate (GSL) levels and profiles of GSLs derived from aromatic amino acids in *Planodes virginica*, *Nasturtium officinale* and *Rorippa sylvestris*.[a]

	N	Total	SD	PE	BAR	EBAR	4hEBAR	IM	4moIM
				105	**40*S***	**40*R***	**139*R***	**43**	**48**
					µmol/g dry wt.				
Planodes virginica (Acc. 1, Kentucky)									
Roots	3	6.8	1.8	0.06	0.05	4.3	n.d.	1.2	n.d.
Stems	3	8.4	3.3	0.001	0.09	6.5	n.d.	0.62	n.d.
Leaves	3	15	4	0.001	0.18	13	tr.	1.4	n.d.
Siliques	3	14	5	0.002	0.17	11	n.d.	0.65	n.d.
Seeds	1	66	–	0.14	0.73	44	0.25	0.11	n.d.
Planodes virginica (Acc. 2, Ohio)									
Leaves	5	24	3	0.002	0.34	17	(0.03)[b]	3.7	n.d.
Nasturtium officinale (two accessions, 'Wild Ohio' and 'Johnny's Select Seeds')									
Seeds (wild Ohio)	1	47	–	31	0.01*	0.07	n.d.	tr.*	tr.*
Seeds (Johnny's)	1	78	–	58	0.01*	0.18	n.d.	tr.*	tr.*
Rorippa sylvestris									
Seeds	2	71	26	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.

N is number of replicates, SD is standard deviation, other data are mean levels.

[a] n-HomoMet-derived and other aliphatic GSLs as specified in Table 5. Searched for but not detected (all three species): The typical *Barbarea* seed isoferuloyl derivatives (6'iF PE, 6'iF IM, 6'iF BAR, 6'iF EBAR), as well as hypothetic isoferuloyl and benzoyl derivatives of other dominating GSLs. Searched for but not detected in *N. officinale*: BZ, 4hBZ, 4hEBAR. Searched for but not detected in *P. virginica* and *R. sylvestris*: all substituted IMs: 4hIM, 1moIM, 4moIM, 1,4moIM. Searched for and not detected in *P. virginica*: BZ, 4hBAR, hydroxy-methoxy derivatives 4moEBAR and 3/4 hmoEBAR ("**x1**"). GSL data with an asterisk indicate lack of MS2 confirmation of analyte identity due to low level.

[b] Confirmed by UHPLC-QToF (Supplementary Fig. S4) and quantified tentatively based on $N = 3$ samples.

Table 5
Profiles of glucosinolates (GSLs) derived from aliphatic amino acids in various organs of *Planodes virginica*, *Nasturtium officinale* and *Rorippa sylvestris*.[a]

	Methylsulfinylalkyls (MeSOalkyls)						$MeSO_2$alkyls	Methylthioalkyls (MeSalkyls)				
	n = 4	n = 5	n = 6	n = 7	n = 8	n = 9*	n = 8	n = 5*	n = 6*	n = 7	n = 8	n = 9*
	64	**72**	**67**	**66**	**69**	**68***	**80**	**94***	**88***	**87**	**92**	**[89]***
							µmol/g dry wt.					
Planodes virginica (Acc. 1, Kentucky)												
Roots	tr.	0.04	0.02	0.28	0.79	0.05	0.01*	n.d.	tr.*	0.02	0.02	n.d.
Stems	tr.	0.01	0.04	0.28	0.78	0.05	0.01*	n.d.	n.d.	n.d.	n.d.	n.d.
Leaves	tr.	0.01	0.02	0.18	0.48	0.05	0.01*	n.d.	n.d.	n.d.	n.d.	n.d.
Siliques	tr.	0.02	0.04	0.43	1.2	0.09	0.03*	n.d.	n.d.	n.d.	n.d.	n.d.
Seeds	tr.[b]	0.10[b]	0.38	4.8	13	0.48	0.05*	tr.	tr.*	0.48	1.3	tr.*
Planodes virginica (Acc. 2, Ohio)												
Leaves	tr.[b]	0.04	0.20	1.2	1.7	0.11	0.02*	n.d.	n.d.	n.d.	n.d.	n.d.
Nasturtium officinale												
Seeds (OH)	tr.*	0.02	0.09	3.9	3.9	0.04	0.17	n.d.	0.8	4.8	2.0	n.d.
Seeds (John)	tr.*	0.02[b]	0.12	4.9	5.5	0.12	0.12	n.d.	1.5	5.1	2.5	n.d.
Rorippa sylvestris												
Seeds	n.d.	n.d.	n.d.	0.06	56	3.2	1.1	n.d.	n.d.	0.02*	10	0.05*

[a] Data (mean levels) and *N* from the same collection as in Table 4. In addition to the listed GSLs, an unidentified isomer of 2h3mPe (**149**), with the desulfo derivative closely eluting yet chromatographically distinct from d**149**, was detected at ca. 0.01 µmol/g dry wt. in *P. virginica* (both accessions), and unidentified isomers of 2mBu and 3mPe at trace level were detected in *N. officinale* (Section 2.2.4.). Searched for but not detected in any of the three species: 10mSOd, 10mSd, Pren, Buen, the typical *A. thaliana* GSLs 3hPr, 4hBu, 3BzOp, 4BzOb, 5BzOp, as well as BCAA-derived 1mEt, 1mPr, 2mPr, 2mBu, 2h2mBu, 3mPe, 3hmPe, 2h3mPe. GSL numbers labelled with an asterisk indicate tentative identification due to lack of an authentic standard, while individual data with an asterisk indicate lack of MS2 confirmation of analyte identity due to low level.

[b] Identity confirmed by comp. witn an authentic std. using UHPLC-QToF (Supplementary Fig. S4F-G).

S5). Occasional reports of BZ and 4hBZ in the species had been challenged before (Agerbirk et al., 2014), again these GSLs were not detected (Table 5). Interestingly, the GSL profile closely matched that of *P. virginica* in a qualitative sense, while several quantitative differences were noticed (Table 5). Most obvious was the dominance of PE over EBAR in watercress and an opposite pattern in *P. virginica*. In both species, EBAR dominated over BAR. A search for both the known isoferuloyl derivative of PE and hypothetic derivatives of other main GSLs did not detect these species.

GSLs derived from BCAAs without chain elongation were searched for but not detected in watercress seeds. Likewise, GSLs derived from 1-2homoIle were searched for but not detected. However, isomers of 2mBu and 3mPe were detected at minute levels using the ion trap HPLC-MS (Supplementary Fig. S5C). Correct high-resolution mass and clear baseline separation from desulfated 2mBu and 3mPe was confirmed by UHPLC-QToF-MS/MS (Supplementary Fig. S5D). Interestingly, the peaks exactly matched minute peaks from *A. thaliana* Col-0 dGSLs (results not shown). As the corresponding 1-2homoLeu derived GSLs have been suggested (but never evidenced) from *A. thaliana* (e.g. Knill et al., 2008) and suggested as products of *A. thaliana* biosynthetic enzymes expressed heterologously (Wang et al., 2020), the minute peaks in *N. officinale* may be the so far hypothetic "3mBu" and "4mPe" derived from 1-2homoLeu. These would be new for the tribe but NMR-spectral evidence for their existence in any plant is still missing (Blažević et al., 2020).

2.2.5. Horseradish

The second crop in the tribe is horseradish (*A. rusticana*). A recent dedicated report discussed the complex GSL profile of horseradish (Agneta et al., 2014), but the report gave many tentative identifications, including a suggestion of either BAR or EBAR in the species at levels around 10–20% of the level of the very low level constituents Buen, Peen and 4moIM. As this suggestion was specifically relevant to our line of inquiry, we aimed to clarify which isomer was present (either BAR or EBAR) but did not detect any isomer at all. From this preliminary investigation, we detected many other GSLs with certainty (Fig. 4).

Samples of horseradish leaf petioles and leaf laminae differed somewhat in the profile of minor GSLs. In samples of the lamina (Fig. 4A), minor peaks included 3mSOp (**73**), the biosynthetic precursor of dominating Pren, as well as Buen, Peen, BZ, IM and 4mIM, all detected well above the noise level but 2–3 orders of magnitude lower than the dominating peak of Pren (5 μmol/g fr. wt.). This result suggested that our sensitivity was just sufficient for trace level detection of BAR or EBAR at the reported level. However, neither BAR nor EBAR were detected. Two branched side chain GSLs tentatively identified by Agneta et al. (2014), 1mPr (**61**) (0.1–0.5 μmol/g fr. wt. in our material) and 2mPr (**62**) (0.02–0.1 μmol/g fr. wt.), were also confirmed in two accessions of horseradish leaf laminae and petioles by t_R, MS and MS2 (both dominated by the fragment $[thioGlc+Na]^+$ as for the authentic standards). Tentatively identified 3mSb (see insert in Fig. 4A), deduced from reasonable t_R, high resolution MS and MS2, agreed with a report by Dekić et al. (2017) and was completely expected as the biosynthetic precursor of 3mSOp and Pren. The observed peak was definitely not the desulfo derivative of the tentatively suggested isomer 4-mercaptobutylGSL (HSbu) (Agneta et al., 2014), as that isomer at our conditions would be accompanied by the oxidized dimer, desulfo $(Sbu)_2$, which was not detected.

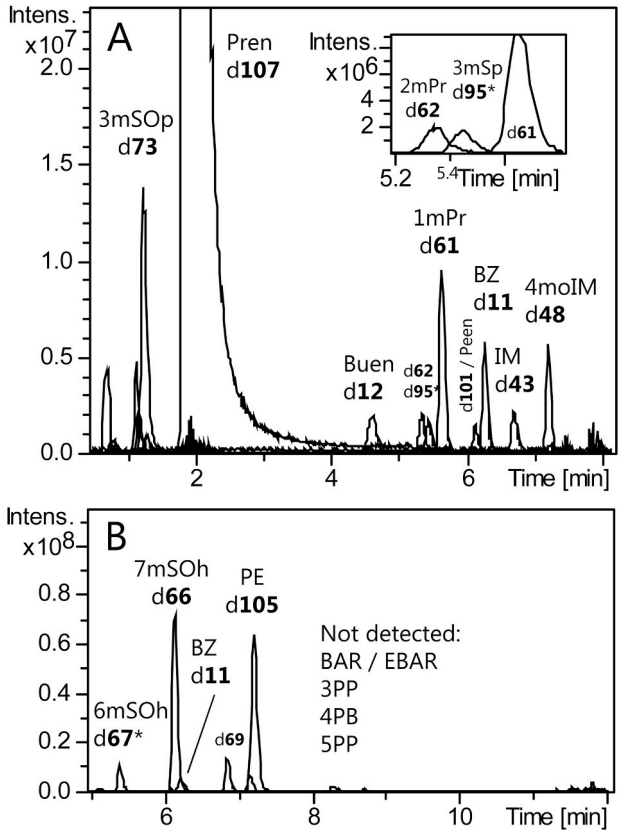


Fig. 4. HPLC-MS chromatogram of desulfoglucosinolates (dGSLs) prepared from glucosinolates (GSLs) from leaves (A) and petioles (B) of horseradish (*Armoracia rusticana*), focusing on trace level GSLs. The chromatograms were made by combining extracted ion chromatograms corresponding to $[M+Na]^+$ of the indicated dGSLs, from analyses that were much overloaded with respect to the dominating dGSL d**107** from Pren. In panel A, an insert shows magnification of the chromatogram from 5.2 to 5.8 min. Neither suggested BAR nor EBAR were detectable. In panel B, only extracted ion chromatograms of *m/z* 382 (BAR/EBAR), 352 (BZ), 366 (PE), 380 (3PP), 394 (4PB), 408 (5PP at high t_R and 6mSOh at 5.4 min), 422 (7mSOh) and 436 (8mSOo) are included. Unlabeled trace peaks did not exhibit a combination of t_R and *m/z* suitable for any of the mentioned candidates. HPLC conditions as in Olsen et al. (2016). Panel A depicts analysis of the Copenhagen garden accession; panel B from the naturalized population at Lake Furesö. An asterisk after a peak number indicates tentative identification.

In the horseradish leaf laminae discussed above, with lack of detected BAR and EBAR, even the frequently reported PE was not detected. When PE was absent, we were possibly not investigating a suitable plant part for BAR or EBAR. However, in petioles we detected PE conclusively by MS2 (accompanied by the same dominating Pren as in leaf plates). Even in petioles, neither BAR nor EBAR were detected (Fig. 4B). We also searched petioles for the recently suggested higher homologs 3PP (**[106]**), 4PB (**[104]**) and 5PP (**[156]**), tentatively suggested natural GSLs so far only deduced from minute peaks of the isothiocyanates from horseradish (Dekić et al., 2017). No candidate peak of any of these homologs was detectable, although comparison of t_R for BZ and PE would suggest them to be detectable at our conditions. Hence, we confirmed previously reported PE but could not confirm the presence of further homologs of PE, nor BAR or EBAR, in our Danish material (two accessions). Searching for 5PP, we unexpectedly detected the isobaric 6mSOp (**67**) in petioles, and furthermore the higher homologs 7mSOh (**66**) and 8mSOo (**69**) (Fig. 4B). Despite the inability to detect higher homologs of PE, our findings were generally in agreement with those of Dekić et al. (2017), including long chain methylsulfinyls and 1mPr (**61**). GSLs derived from 1 to 2 homoIle were searched for but not found. Obviously, the GSL profile of horseradish is highly complex and organ specific (Agneta et al., 2014; Dekić et al., 2017) and deserves further investigation.

2.2.6. The emerging model species Cardamine hirsuta

In a continuation of our previous study of *Cardamine* species, we investigated *C. hirsuta*. A few reports of the GSL profile exists (reviewed by Olsen et al., 2016; Bakhtiari et al., 2018), but further inquiry was deemed relevant. Two accessions were used from nearby Copenhagen University gardens but collected from different soil types (Section 4.2.). To our surprise, the accessions exhibited contrasting GSL profiles from each other (Supplementary Fig. S6), while each individual accession was homogeneous among individuals in qualitative GSL profile. Concerning aromatic GSLs (Table 6), both were dominated by BZ, and both accessions showed the ability to form a range of Trp-derived GSLs (IM and derivatives), with a tendency for a different balance of derivatives such as 1moIM (Table 6).

Concerning aliphatic GSLs, the two *C. hirsuta* accessions showed both similarities and contrasts (Table 7). Both contained only Met-derived GSL from at least two chain elongations ("C4 group" and above), and both showed high levels of Buen. However, accession 1 additionally showed high levels of the β-hydroxylated *R*2hBuen, while accession 2 showed high levels of the homolog Peen. All three GSLs were conclusively detected after desulfation by t_R and MS2 comparison with authentic standards identified by NMR (Fig. 5, Supplementary Fig. S6). In particular, the MS2 of desulfo Peen is remarkably complex and useful as fingerprint (Fig. 5) due to the fragment types j and k (Table 1). Possibly, the extra methylene group compared to Buen brings the side-chain double bond of Peen in proximity to the SGlc moiety, resulting in combined movement of the Glc residue to the N–OH oxygen and subsequent oxidation to gluconolactone by the double bond.

The polymorphism in chain elongation is reminiscent of the well-known C3–C4 polymorphism in *A. thaliana* (Kliebenstein et al., 2001) and a general distinction of C3 and C4 species (Windsor et al., 2005; Agerbirk et al., 2008), but in this case between C4 and C5, suggesting

Table 6
Total glucosinolate (GSL) levels and profiles of GSLs derived from aromatic amino acids in two accessions of *Cardamine hirsuta*.[a]

	N	Total	SD	BZ	PE	IM	1moIM	4hIM	4moIM
				11	**105**	**43**	**47**	**28**	**48**
Organ						µmol/g dry wt.			
Accession 1 (Skyggehöj) on peat-based soil among cultivated *Rhododendron* spp.									
Roots	1[b]	26	–	20	n.d.	0.88	0.01*	0.06	0.51
Leaves	4[c]	49	12	19	0.04*	0.06	n.d.	0.31	0.19
Siliques	1[d]	16	–	8.3	0.02*	n.d.	n.d.	n.d.	n.d.
Accession 2 (Villa Rolighed) on loamy, clayey soil.									
Roots	1[b]	41	–	29	1.3	0.92	0.99	0.02*	0.41
Leaves	4[c]	50	6	19	0.5	2.5	0.02	0.01*	0.07
Siliques	1[d]	27	–	6.7	0.3	0.78	n.d.	n.d.	n.d.

N is number of replicates, SD is standard deviation, other data are mean levels.
[a] n-HomoMet-derived and other aliphatic GSLs as specified in Table 7. Searched for but not detected: BAR, EBAR, 1,4moIM. Pooled stems ($N = 1 + 1$) showed a similar profile as other plant parts (Acc 1: mean 11 µmol/g dry wt. Acc 2 mean 8.5 µmol/g dry wt.)., except that a possible trace of 4hBZ (**23**) was detected in one sample of stems (Accession 1). GSL data with an asterisk indicate lack of MS2 confirmation of analyte identity due to low level.
[b] Pooled roots from the population.
[c] Mean for one batch of pooled cauline leaves, one batch of pooled rosette leaves and two samples of cauline leaves from individual plants. Samples of individual plants showed the same qualitative profile as seen in pools, and rosette leaves and cauline leaves showed similar profiles.
[d] Pooled siliques from the population.

Table 7
Profiles of glucosinolates derived from aliphatic amino acids in two accessions of *Cardamine hirsuta*.[a]

	4mSOb	Buen	5mSOp	Peen	*R*2hBuen	*S*2hBuen	8mSOo
Global No.	**64**	**12**	**72***	**101**	**24*R***	**24*S****	**69**
Organ				µmol/g dry wt.			
Accession 1 (Skyggehöj) on peat-based soil among cultivated *Rhododendron* spp.							
Roots	0.43	0.6	n.d.	n.d.	2.5	tr.*	0.58
Leaves	0.75	6.5	n.d.	n.d.	21	tr.*	0.26
Siliques	0.09	0.5	n.d.	n.d.	6.8	tr.*	0.14
Accession 2 (Villa Rolighed)) on loamy, clayey soil							
Roots	0.28	3.4	0.6	3.5	0.03	0.02*	0.08
Leaves	0.29	14	1.6	11.7	0.06	0.03*	0.29
Siliques	0.06	12	0.2	7.5	0.08	0.02*	0.12

[a] Data (mean levels) and *N* from the same collection as in Table 6. Pooled stems ($N = 1 + 1$) showed a similar profile as other plant parts (Acc 1: mean 9.4 µmol/g dry wt. Acc 2 mean 14 µmol/g dry wt.). Searched for but not detected: 1mEt, 1mPr, 2mPr, 2mBu, 2h2mBu, 3mPe, 2h3mPe, 3hmPe, Pren, 4mSb, 8mSO$_2$o. GSL numbers labelled with an asterisk indicate tentative identification due to lack of an authentic standard (no auth. std. of **72** was available when this sample was run in 2012), while individual data with an asterisk indicate lack of MS2 confirmation of analyte identity due to low level.

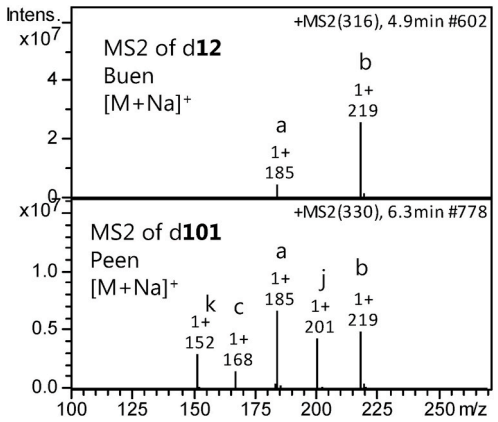


Fig. 5. MS2 spectra of desulfated glucosinolates (GSLs) confirming the identity of two 2-3homoMet-derived GSLs in *C. hirsuta*. While the spectrum of desulfo Buen contains only the usual fragments for this type of dGSL (a, [anhydroGlc+Na]$^+$; b, [thioGlc+Na]$^+$), the spectrum of desulfo Peen contains an additional usual fragment (c, [M-anhydroGlc+Na]$^+$) and two unusual fragment ions suggesting a structure-specific cyclization and exchange of O during fragmentation: 201, [gluconolactone+Na]$^+$ and 152, [$C_6H_{11}NS$+Na]$^+$. The unusual fragments can be rationalized as fragment a plus O and fragment c minus O, respectively.

chain elongation machinery that deviates from the previously studied (Textor et al., 2007; Kumar et al., 2019; Petersen et al., 2019). The polymorphism in hydroxylation at β is likewise similar to a known polymorphism in *A. thaliana* (Hansen et al., 2008). Although we propose that this polymorphism is genetically controlled, this cannot be concluded from this report, as we only sampled natural populations growing in contrasting environments.

Both *C. hirsuta* accessions also contained the expected biosynthetic precursors 4mSOb and/or 5mSOp and the higher homolog 8mSOo. *R*2hBuen was known from the species (Bakhtiari et al., 2018) but has otherwise only once been demonstrated conclusively in the tribe Cardamineae (Olsen et al., 2016). The trace level *S*2hBuen was a new GSL for the tribe, but completely expected as biosynthesis of one β-hydroxy GSL is usually accompanied by the epimer to some degree.

2.2.7. A search for aliphatic glucosinolates in Barbarea vulgaris

As observed for the Australian *Barbarea* species (Section 2.2.1.), aliphatic GSLs generally appear to be absent from the genus, except for a single historical report of isothiocyanate (ITC) hydrolysis products of aliphatic GSLs, prop-2-enyl ITC, isopropyl ITC and 3-(methylthio)propyl ITC, from *Barbarea* spp. (Cole, 1976). However, identification of some ITCs in that report was of questionable reliability (Olsen et al., 2016). Indeed, that historical paper also (unexpectedly) reported prop-2-enyl ITC from *Plantago major* L., (Plantaginaceae: Lamiales] although competent researchers were later unable to detect any GSL in that

species (Larsen et al., 1983). However, more recently data also suggested the possibility of aliphatic GSLs in *Barbarea*. Three observations were particularly suggestive: CYP79F6 and other putative GSL biosynthetic enzymes were reported to be induced by larvae of the moth *Plutella xylostella* (L.) (Liu et al., 2016), *B. vulgaris* amino acid chain elongating enzymes were reported to chain elongate Met and Leu in addition to Phe (Wang et al., in press), and putative GSL products were induced by spraying leaves with 10 mM $CuCl_2$ (*aq.*) (Pedras et al., 2015). Hence, we decided to search for proposed aliphatic GSLs after various kinds of induction and from normal growth conditions. Specifically, we re-investigated both ecotypes of *B. vulgaris* for 2-propenylGSL (Pren, **107**) and related aliphatic GSLs by ion trap HPLC-MS inspection for trace peaks, experimental testing of recovery of added GSLs, and investigation of plants challenged by herbivory and $CuCl_2$. We also investigated the morphologically deviating *B. vulgaris* ssp. *vulgaris*.

The GSL precursor of 2-propenyl ITC, Pren, was not detectable in *B. vulgaris* (Fig. 6B), in agreement with previous investigations of a range of P-type and G-type accessions from a large geographical range or harvested at different seasons (Agerbirk et al., 2015 and references cited therein). Experimental addition of (intact) Pren to crude extracts followed by desulfation confirmed linear detection of this GSL down to trace levels (Fig. 6C–E). The limit of detection was estimated as 0.1 μmol/g dry wt. at normal sample concentration (Fig. 6E) and 10-fold lower when using concentrated samples or high injection volumes. In the latter case, however, linearity of dominating GSLs suffered. The morphologically deviating ssp. *vulgaris* (in this report represented by accession B59), characterized by erect (vertical) rather than arcuate siliques, had a GSL profile indistinguishable from that of G-type plants (results not shown), in agreement with a previous report including ssp. *vulgaris* (Agerbirk et al., 2003).

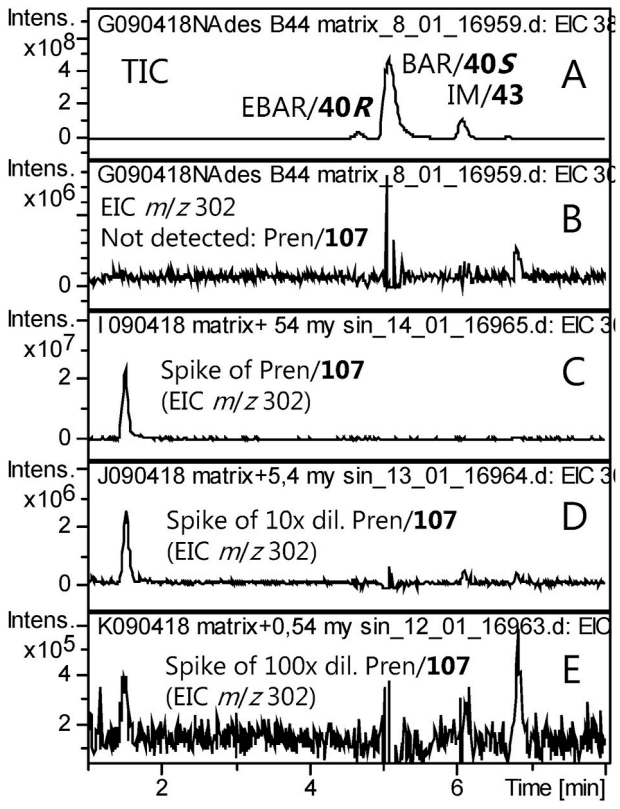


Fig. 6. Lack of the aliphatic glucosinolate (GSL) Pren (107) in *B. vulgaris* and spiking of pure (intact) Pren for establishing the limit of detection. A. Total ion chromatogram for the three dominating peaks in G-type *B. vulgaris* (dGSL preparation). B. Extracted ion trace for desulfo Pren in the same extract as A, showing lack of detection. C, D, E. Results of serial spiking of the crude extract with serial 10-fold dilutions of Pren before the desulfation procedure, showing linearity also at low levels and ability to detect trace levels.

We subjected P and G-type *B. vulgaris* to herbivory by larvae of the butterfly *Pieris brassicae* L. (3 days or 7 days), or by *P. xylostella* larvae (4 days), to spraying with 10 mM $CuCl_2$ (harvest after 4 days), and a parallel control treatment. These treatments were followed by GSL analysis using HPLC-UV for observing the general GSL profile and by ion trap HPLC-MS with injection of quite concentrated samples for a search for induced GSLs. The general leaf GSL profile was as expected for the type in both control and challenged plants, and we observed no candidate induced GSL peak. Some rather drastic treatments resulted in lower mean GSL levels, such as herbivory by the large *P. brassicae* for 7 days, after which time much of the leaves had been eaten (Fig. 7). The lower GSL levels in remaining leaves may well be due to preferential ingestion of GSL rich plant parts of these GSL-stimulated larvae. We carefully searched for but did not detect Pren in any HPLC-MS sample in extracted ion chromatograms. Likewise, we systematically searched for but did not detect 1mEt, 1mPr, 2mPr, 2mBu, 2h2mBu, 3mPe, 2h3mPe, 3hmPe, 4mSb, 4mSOb, Buen, 8mSOo or BZ, selected from prominent GSLs in other tribe Cardamineae members and from GSLs produced by combination of *A. thaliana* and *B. vulgaris* biosynthetic enzymes after heterologous expression in tobacco (Wang et al., in press). Neither did visual inspection reveal any other new GSL for the species. However, we did detect trace levels of the known minor constituents 4mIM and PE, included in each search as a positive control.

During the final editing of the present work, a claim of ITCs corresponding to 3mSOp (**73**) and 4mSObuen (**63**) in *P. xylostella*-induced *B. vulgaris* was published (Hussain et al., 2020). The claim, lacking actual evidense for the identifications as well as levels calculated on a plant weight basis, was based on analysis of G-type *B. vulgaris* of the Hedeland accession also used in our induction experiments, and no other isothiocyanate (such as the expected phenethyl ITC corresponding to PE) was reported (Hussain et al., 2020). Hence, we searched three desulfoGSL analysis files of *P. xylostella*-induced *B. vulgaris*, including both the Hedeland and Suserup accessions, for any signs of d**63** or d**73**. We also searched for d**95** from the expected biosynthetic precursor **95**. No trace of any of these dGSLs was detected, although we are able to detect them in general (Table 2). We also searched non-induced control samples with the same negative result. We notice that one of the claimed ITCs (from **73**) would be isobaric with the expected ITC from PE (**105**), and that the corresponding dGSL (d**105**) was detected as a minor constituent in several of our *P. xylostella* induced and control samples.

In conclusion, there are no reliable reports of aliphatic GSLs in the genus *Barbarea*. In contrast, extensive scrutiny has failed to reveal aliphatic GSLs above an estimated limit of detection of 0.1–0.01 μmol/g dry wt.

2.3. Phylogenetic analysis of Barbarea vulgaris accessions

In order to better understand whether the tested types and subspecies of *B. vulgaris* were representative for the entire species, we sequenced the nuclear ribosomal DNA internal transcribed spacer (ITS) region of a range of our accessions. For the P-type, we included our Danish "type-accession" (B4) as well as three accessions representing major parts of northern Russia (Supplementary Table S1), and contrasted with two different morphological types of plants with glabrous leaves: the G-type (Danish "type accession" B44) of ssp. *arcuata* (characterized by arcuate siliques) and a newly collected accession (B59) with erect (vertical) siliques traditionally named ssp. *vulgaris* by Danish botanists (Hartvig et al., 2015).

The sequence variation of our accessions supported distinction of P- and G-types, with only two sequence types detected in the six sequenced accessions. The G-type as well as accession B59 of the morphological type traditionally known as ssp. *vulgaris* shared the same ITS sequence, while all four sequenced P-type accession differed from the G-type sequence at eight nucleotide positions (Table 8) but where identical to each other.

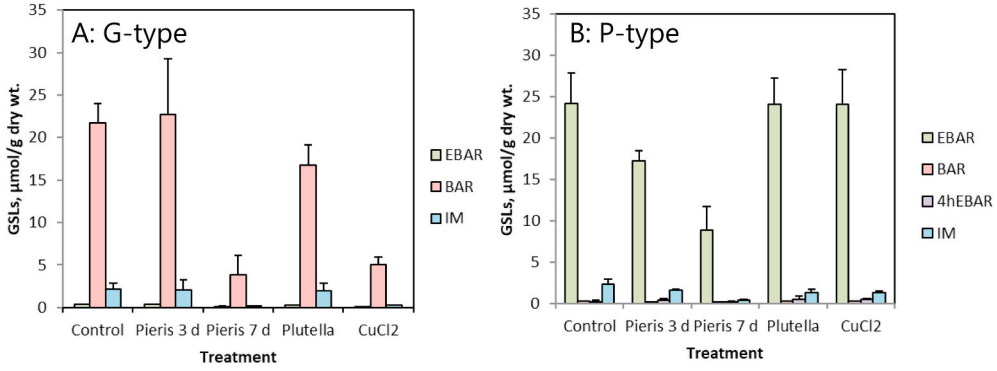


Fig. 7. Levels of major glucosinolates in leaves of first year rosette plants of the G-type (A) and P-type (B) of *Barbarea vulgaris* in plants subjected to various challenges or no challenge as control. The contrasting general profile of the types is evident from dominance of BAR in the G-type and EBAR in the P-type. Treatment codes are: Control, un-challenged plants harvested after 7 days; Pieris 3d and Pieris 7d, herbivory by *Pieris brassicae* larvae until harvest at either day 3 or day 7; Plutella, herbivory by *Plutella xylostella* for 4 days; CuCl2, spraying of leaves with 10 mM $CuCl_2$ (*aq.*) followed by recovery for 4 days. Bars represent means, whiskers indicate standard deviation (N = 3 for each group).

Table 8
Varying single nucleotide positions among ITS sequence variants within *Barbarea vulgaris* including newly obtained and publicly available sequences by mid-2019 (Fig. 8). In total seven groups of sequence types were found, among which group 2 and 3 differ from the G-type only by additional sequence ambiguities not indicated here. Eight sites are distinguishing between P- and G-types as indicated in bold.

Sequence group No.[a]	Number of accessions	Type or subspecies	Position in alignment[b]								
			81	83	181	251	452	474	565	586	617
6	31	P-type[c]	C	A	G	C	C	A	T	T	C
1	21	G-type[c]	T	T	A	T	C	G	C	C	T
7	9	"*rivularis*"[c]	T	T	A	T	T	G	C	T	T
4	3	(PxG?)[d]	C	T	A	Y[e]	C	G	T	T	T
5	2	–	C	T	G	T	T	G	C	T	T
2	1	–	T	T	A	T	C	G	C	C	T
3	1	–	T	T	A	T	C	G	C	C	T

[a] See Supplementary Table S1 and Fig. 8. GenBank accession codes for the newly acquired sequences are: Group 6 (P-type), MN508440; Group 1 (G-type and ssp. vulgaris), MN508439.
[b] Refer to the alignment in Supplementary Table S2.
[c] The designations of P- and G-type represents the majority of G-type and P-type accessions in Supplementary Table S1. Sequences from material identified as ssp. *rivularis* or ssp. *vulgaris* tended to cluster in sequence group 7, although a few ssp. *vulgaris* sequences were also found in sequence group 1.
[d] Preliminary interpretation of the intermediacy of the sequence: DNA sequence information and phylogenetic position in the network of accessions from group 4 might indicate that these accessions represent hybrids between the two clearly separated clades found within *B. vulgaris* according to the SplitTree analysis.
[e] Y: sequence ambiguity, indicating multiple ITS copies within a single individual with differing nucleotides (C or T).

This ITS sequence variation was subsequently compared with all *B. vulgaris* ITS sequences available from GenBank by mid-2019, most of them from Lange et al. (in review), resulting in a slightly more complex pattern. In this and the following discussion, we accepted the original sequence-contributors' choice of intraspecific nomenclature. Currently, no intraspecific taxa are generally accepted and the intraspecific nomenclature of *B. vulgaris* is in need of revision (Lange et al., in review). The names represented with the GenBank material were: G-type, P-type and two traditionally recognized subspecies with more or less erect siliques; ssp. *vulgaris* and ssp. *rivularis* (Hegi, 1958).

In total, seven ITS sequence variants were represented in GenBank (Suppl. Table S1, Table 8). The majority of these sequences likewise grouped with either our G-type or P-types in agreement with their respective type identities. One group (Group 4) was intermediary of the P-type and G-type sequences, and was tentatively interpreted as a hybrid of the G- and P-types. Indeed, one of them had been deposited as a hybrid sequence. Another group (group 7) was dominated by sequences deposited as ssp. *rivularis* and ssp. *vulgaris* and was more similar to the G-type than to the P-type (Table 8), for simplicity this group is named the *rivularis* group here, although the name may be temporary.

A few GenBank sequences did not agree with the group designations used here. Three sequences attributed to PxG hybrids clustered with the G-types in sequence group 1, as did two sequences attributed to ssp. *vulgaris*, including our accession B59. Likewise, one sequence attributed to G-type and two sequences attributed to hybrids grouped with the P-types in sequence group 6. Three other sequence groups (Groups 2, 3 and 5) represented sequences deposited as G-type, GxP hybrid or undetermined. These groups were represented by few deposited sequences.

The parsimony network analysis (Fig. 8) suggested two main clusters of *B. vulgaris*, the one including the P-type and the small group 5 and the other including most of the remaining groups (G-type, *rivularis* and the small sequence groups 2 and 3). Group 4, tentatively interpreted as hybrid of P- and G-types, was intermediate. A separation of *B. vulgaris* into two gene pools, correlated with G-type and P-type, has also been shown earlier using different genetic marker systems (Christensen et al., 2014; Toneatto et al., 2012).

In conclusion, the investigated G-type and P-type of *B. vulgaris* are rather representative for the known genetic variation of the species, which has mainly been investigated in the European part of its main original distribution area (Christensen et al., 2014; Toneatto et al., 2010, 2012). Hence, lack of aliphatic GSLs may be a general property of the species (and genus). However, other sequence groups were found, with the "*rivularis*" group most deviating. These groups may represent other chemotypes as well. Hence, lack of aliphatic GSLs in the genus (Section 2.2.7.) can of course only be strictly concluded for the accessions investigated, not in general for the species and genus.

2.4. Glucosinolates in a distantly related genus, Reseda

To provide a perspective of the diversity in the tribe Cardamineae, the species *Reseda luteola* (family: Resedaceae) was investigated because the literature suggested the GSL profile to remarkably resemble that of *B. vulgaris*. The presence of BAR in *R. luteola* is well-established

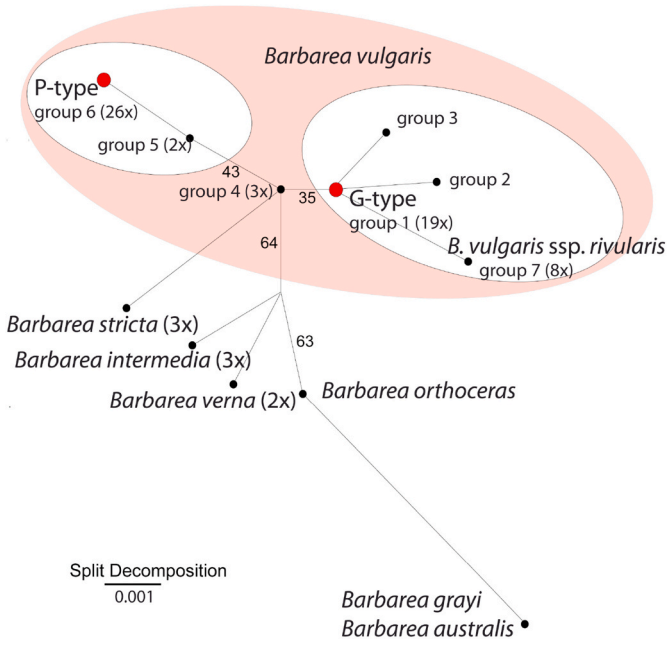


Fig. 8. The intraspecific diversity of *Barbarea vulgaris* and relations to other species in the genus analyzed by parsimony network analysis of ITS regions from *Barbarea* accessions using SplitsTree (Huson and Briant, 2006). The *B. vulgaris* ITS sequences are defined as seven groups and detailed accession information is found in Supplementary Table S1. The respective alignment is shown in Supplementary Table S2, and bootstrap values are provided for 1000 replicates.

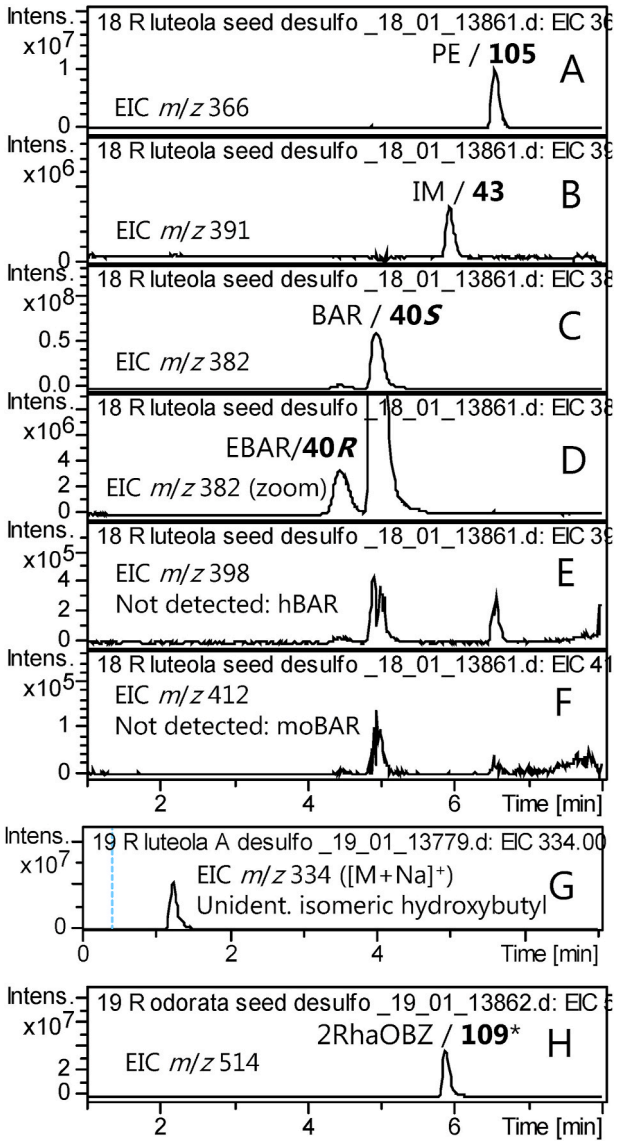


Fig. 9. Extracted ion HPLC-MS chromatograms of desulfoglucosinolates prepared from glucosinolates (GSLs) in seeds (A–F) or leaves (G) of *Reseda luteola* and seeds of *Reseda odorata* (H). The three major peaks (A, B, C) represent PE, IM and BAR, much like many *Barbarea* spp. Focus on minor peaks (D) allowed conclusive identification of EBAR, confirmed by t_R and the characteristic MS2 spectrum. A range of putative derivatives were not detected (E, F), but an unidentified hydroxybutylGSL was present (G), as was a known glycoside in *R. odorata*.

(Kjær and Gmelin, 1958; Bennett et al., 2004), despite the great evolutionary distance to *Barbarea*. We revisited this species in order to check for traces of the epimer EBAR and related GSLs. The major peaks were confirmed to be BAR, PE and IM, in seeds (Fig. 9A–C) and in leaves. A minor peak was conclusively identified as EBAR (Fig. 9D). However, hydroxy or methoxy substituted glucobarbarins were not detected (Fig. 9E–F). As these are available as references from *Barbarea* and *Arabis* spp. and known to be detectable and base line separated from the *R. luteola* peaks at our conditions (Table 2; Olsen et al., 2016), their absence (above the limit of detection) could be concluded with certainty.

The level of EBAR in *R. luteola* seeds corresponded to ca. 5% of the sum of EBAR and BAR as estimated from peak intensities, but was relatively lower in leaves, ca. 1% of the sum of EBAR and BAR. The varying proportion of the epimeric glucobarbarins in seeds and leaves of *R. luteola* suggests that each epimer is due to a distinct genetic locus and biosynthetic enzyme, as is the case in *B. vulgaris* (Liu et al., 2019a).

Putative constituents like BZ, 1mEt, 1mPr/2mPr and isomers of mBu and mPe were searched for in both leaves and seeds of *R. luteola* but not detected (results not shown). However, an apparent isomer of hydroxybutylGSL was detected (Fig. 9G), suggesting that use of amino acid precursors for seed GSL biosynthesis in *R. luteola* is not quite as restricted, to Trp and homoPhe, as in *B. vulgaris*, but include an aliphatic precursor.

We further investigated the apparent hydroxybutylGSL in *R. luteola*. Only trace-levels were found in seeds, but analysis of leaves revealed appreciable levels of the apparent hydroxybutylGSL (Fig. 9G), which eluted at 1.3 min and hence was not 1hmPr (**30**) but could be either Met-derived 4hBu (**[26]**) (not previously known from basal families) or Leu-derived 2h2mPr (**31**), not available in our reference library at that time, or several yet unknown isomers. Ion trap MS2 was inconspicious and identical to both the MS2 of d **[26]** and d**30**. We re-analyzed the sample using UHPLC-QToF MS/MS and found the high-resolution mass to be in agreement with a desulfo hydroxybutylGSL (found: 312.1114 (mean, N = 2), calc. for $C_{11}H_{22}O_7NS^+$([M+H]$^+$): 312.1112). This result was in agreement with a previous report of an unspecified hydroxybutyl GSL in leaves of *R. luteola* (Griffiths et al., 2001).

Neither BAR nor EBAR were detected in *Reseda odorata* seeds in the present study, in contrast to a previous report (Bennett et al., 2004). The dominating dGSL peak at *m/z* = 514 was presumably due to the well-established 2RhaOBZ (**109**) (Pagnotta et al., 2020). The dGSL d**109** showed a distinct MS2 compared to that from the isomer d**110** (Supplementary Fig. S1C) and confirmed our ability to detect a wide range of structural types including glycosides (Fig. 9H). A minor peak of desulfo IM almost coeluted with desulfo 2RhaOBZ. A range of aliphatic GSLs including butyls and hydroxybutyls were searched for but not detected in seeds.

While both *Reseda* species accumulated IM, neither of the usual IM derivatives for tribe Cardamineae (Supplementary Fig. S1A) were

detected, although they were specifically searched for. No previous author has reported IM derivatives from the genus, even when roots were examined (Pagnotta et al., 2020). However, roots should be investigated critically before concluding on presence or absence of substituted indole GSLs in this genus.

3. Conclusions

Aided by a library of t_R and MS/MS of 75 generally accepted GSLs, we re-investigated selected species from the tribe Cardamineae and confirmed presence of numerous GSLs. The increased analytical sensitivity due to HPLC-MS relative to HPLC-UV tended to increase the number of detected GSLs in the species. Even when using quite sensitive methods, GSL profiles of individual species showed distinct qualitative differences from each other, suggesting that individual species and evolutionary lines exhibit distinct qualitative similarities and differences. This finding is in contrast to the informal saying "everything is everywhere, detection or not is a matter of analytical sensitivity", which may apply for elemental analysis but apparently not for GSL analysis. The distinctness of the GSL profiles suggests that GSL profiles reflect presence or absence of biosynthetic activities with characteristic specificities. For example, the diversity of minor GSLs and deduced precursor amino acids was astounding in horseradish (Fig. 4), and contrasted with the lack of detection of similar minor GSLs in *B. vulgaris* (Fig. 6). We also observed contrasting chain length profiles between species (*R. sylvestris* vs *P. virginica* and *N. officinale*) and within one species (*C. hirsuta*).

Acylation at the thioglucose moiety was solely detected in the genus *Barbarea* despite careful search in relevant organs (seeds) of other species. Likewise, aliphatic GSLs were not detected in *B. vulgaris* despite extensive investigations. Similarly, the rare δ-hydroxylated GSL 3hmPe (**141**) and the β-hydroxylated isomer 2h3mPe (**149**) were not detected in any of the investigated species and seem so far to remain unique for *C. pratensis*. In contrast, many other GSLs were more wide-spread than previously anticipated, such as BCAA-derived GSL confirmed in horseradish and tentatively confirmed in watercress. The epimeric glucobarbarins (BAR and EBAR) were quite widespread in the tribe (confirmed in *P. virginica*, watercress and all *Barbarea* spp), but were not detected in horseradish. A phenolic derivative (4hEBAR, **139*R***) previously believed to be unique to *Barbarea* was also detected in *P. virginica*, albeit at very low levels, but not in watercress or *R. luteola*. Evolutionary patterns in the data are discussed in the accompanying paper (Agerbirk et al., 2021).

Despite the large battery of references, additional GSLs were occasionally detected, such as an unidentified possibly BCAA-derived GSL in watercress and *A. thaliana*. Comparison of t_R was repeatedly critical in peak identification, and we conclude that authentic references are essential for reliable GSL analysis. Two new reference materials (*P. virginica* and watercress) were added to the existing range available to science, as was a library of systematically assigned MS2 spectra.

Two cases of intraspecific polymorphism in GSL profiles were investigated. The very well investigated P- and G-types of *B. vulgaris* were well defined and quite well separated in split-tree analysis based on ITS sequences. The G-type seemed to be representative for *B. vulgaris* often studied, and the apparent lack of aliphatic GSLs in the species (above the limit of detection) was tentatively concluded to be typical for the species in general. A hitherto unknown qualitative polymorphism in *C. hirsuta* was also discovered, adding to the characters that can be investigated in this emerging model plant. Such documented GSL polymorphism appears to be rare in the tribe and has evolutionary implications if genetically controlled (Agerbirk et al., 2021).

4. Experimental

4.1. General experimental procedures

Potassium (3*E*)-4-(methylthio)but-3-enylGSL (**83**) and potassium (R_S, 3*E*)-4-(methylsulfinyl)but-3-enylGSL (**63**) were authentic standards from Centro di Recerca per le Colture Industriali, Bologna, Italy. Potassium prop-2-enylGSL (Pren, **107**) and potassium 3-(methylsulfinyl) propylGSL (3mSOp, **73**) were authentic standards from Carl Roth, Karlsruhe, Germany. Desulfated 3mSOp agreed with the major peak in desulfoGSLs from *Iberis amara* seeds. All intact reference GSLs were converted to dGSLs as usually. A minor peak in the same desulfated seed extract agreed with desulfo 3mSp (d**95**) with respect to mass (found, 328.0890 (mean, N = 6), calc. for $[C_{11}H_{22}O_6NS_2]^+$: 328.0883), MS and MS2 in UHPLC-QToF, this peak was used as reference for ion trap HPLC-MS. Other authentic references were obtained as listed (Table 2). For induction of plants, $CuCl_2$ dihydrate (>99.99%) was from Aldrich (now Sigma-Aldrich, St. Louis, MO). For HPLC, acetonitrile ("≥99.95%, HiPerSolv CHROMANORM® for LC-MS") was from VWR, Søborg, DK, while HCOOH ("LC-MS Ultra") was from Fluka, Roskilde, DK.

4.2. Plants and insects

As "type" accessions of the G-type and P-type of *Barbarea vulgaris* W. T. Aiton (Brassicaceae), we used accessions also included in our initial characterization (Agerbirk et al., 2003) and later subjected to transcriptome analysis and genetic mapping (Liu et al., 2016, 2019a): seeds of the G-type of *Barbarea vulgaris* were of accession B44 (NGB31789) originating from the same population as originally numbered B1 (Agerbirk et al., 2003) while seeds of the P-type of the same species were of accession B4 (NGB23547) (Agerbirk et al., 2003). Both accessions agree with descriptions of ssp. *arcuata* (Opiz.) Fries in standard floras (e.g. Hegi, 1958), but are currently not named to subspecies level due to the biological and phylogenetic variation discussed here and elsewhere. Both accessions are available from https://nordic-baltic-genebanks.org, registered under the indicated NGB numbers.

A new *Barbarea vulgaris* accession was *B. vulgaris* ssp. *vulgaris* (with vertical siliques in contrast to arcuate siliques of var. *arcuata*) collected June 6, 2018, as plants by NA at a disturbed roadside/garden edge at decimal coordinates 56.083, 12.516 (Stormlugen 9, 3140 Ålsgårde, DK), given the N. Agerbirk accession number B59, cultivated 2018–2019 in Kgs. Lyngby (DK) (consistently showing vertical siliques) and deposited at NordGen as NGB31971.

Russian *Barbarea vulgaris* P-type accessions from a previous report (Christensen et al., 2014) were multiplied in isolation from any other flowering *Barbarea* spp. and given N. Agerbirk accession numbers and NordGen numbers as follows: Mix of Mosc1+2+3 propagated 2015–2016 in Kgs. Lyngby (DK): B54 (NGB 31598). Komi propagated 2015–2016 in Aarhus (DK): B55 (NGB 31590). Whit propagated 2016–2017 in Kgs. Lyngby (DK): B56 (NGB 31790).

Several vouchers of both first and second year *Barbarea vulgaris* plants (B4, B44, B56, B59) have been deposited at CP (Supplementary Fig. S7).

Barbarea grayi Hewson (Brassicaceae) was obtained as seeds in 2013 from Royal Botanic Gardens, Melbourne.

Barbarea australis Hook f. (Brassicaceae) was obtained as seeds from Threatened Species Unit, Dept. of Primary Industries, Water and Environment, Hobart, Tasmania, Australia, as previously reported (Agerbirk and Olsen 2011). Material was propagated 2002–2003 in Kgs. Lyngby (DK) and is available at NordGen (https://nordic-baltic-genebanks.org) as accession NGB23039. A voucher (rosette) has been deposited at CP (Supplementary Fig. S7).

Cardamine hirsuta L. (Brassicaceae) was represented by two accessions, both from the university gardens at University of Copenhagen, Frederiksberg Campus, and identified by MØ. Accession 1 was collected June 21, 2012 by NA, growing as a weed in moist, peat-based soil among cultivated *Rhododendron* spp. ("Skyggehøjen", Landbohøjskolens Have), while accession 2 was collected June 25, 2012 by MØ and NA, growing as a weed in a loamy soil at "Villa Rolighed". A voucher of each accession has been deposited at CP (Supplementary Fig. S7).

Planodes virginica (L.) Greene (Brassicaceae) (The synonym

P. virginicum (L.) Greene is preferred by worldfloraonline)was represented by two accessions both identified and collected by DC. Accession 1 ("Kentucky") was collected as seeds on the campus of Northern Kentucky University. Individual plants were grown from seeds. Three of them were harvested, divided in roots, stems, leaves, and siliques and dried. A voucher was deposited at the Wright State University Herbarium (Supplementary Fig. S7). Accession 2 ("Ohio") was collected as leaves of wild growing plants early in the season in 2015 at Cox Arboretum, Dayton (OH) and dried. Seeds and dry samples were sent by mail to the Copenhagen lab for analysis. The seeds were propagated by NA and deposited at NordGen as accession NGB31591.

Rorippa sylvestris (L.) Besser (Brassicaceae) with (sparse) seeds was collected October 2017 and identified by NA at the gardens of Dept. of Plant and Environmental Sciences, Univ. of Copenhagen. Leaf GSL profile was also checked and found similar to profiles reported by Olsen et al. (2016). A voucher of the same species was previously deposited at CP (Olsen et al., 2016).

Nasturtium officinale W.T. Aiton (Brassicaceae) (watercress) seeds were of two accessions. Accession 1 was identified and collected by DC from a wild population from Yellow Springs (OH). Accession 2 consisted of commercial seeds obtained from Johnny's Selected Seeds (www.johnnyseeds.com).

Armoracia rusticana P. Gaertn., B. Mey. & Scherb. (Brassicaceae) (horseradish) was identified and collected by NA. For extraction and GSL analysis we used fresh leaves from a naturalized population (Lake Furesø, Denmark) and from a clone of unknown origin propagated in a Copenhagen garden. A voucher of vegetative parts of this clone, that did not flower as is usual for the species, has been deposited at CP (Supplementary Fig. S7).

Seeds of *Reseda luteola* L. (product number 1086N), *Reseda odorata* L. 'grandiflora' (product number 1087D) and *Iberis amara* L. (product number 703E) (all Brassicaceae) were from Chiltern Seeds, Wallingford, UK. Leaves of *R. luteola* were from the same species as cultivated in the gardens of Dept. of Plant and Environmental Sciences. Seeds of *Arabidopsis thaliana* (L.) Heynh. Col-0 (Brassicaceae) were as described elsewhere (Graeff et al., 2016).

Insect and general insect-experimental conditions were as reported elsewhere (Christensen et al., 2019). For the experiments with insect- and $CuCl_2$-challenged *B. vulgaris* plants, P-type accessions were collected as seeds in Denmark at either Trundholm (two plants per treatment) or Brokøb, while G-type accessions were from either Suserup (two plants per treatment) or Hedeland (Heimes et al., 2016). For experiments with $CuCl_2$-challenged plants, leaves were sprayed below and above with 10 mM $CuCl_2$ (aq.) using a domestic spray flask.

4.3. Glucosinolate analysis by HPLC-UV and ion trap HPLC-MS/MS

In general, plant material was freeze dried before analysis, by placing the relevant material directly in the freeze drier without previous freezing and leaving it until crisp (Agerbirk et al., 2014). However, seeds stored at ambient humidity were considered dry and extracted directly (reported levels were not corrected for any residual moisture), and *B. vulgaris* foliage from the induction experiment was dried over silica gel until crisp rather than freeze-dried (Christensen et al., 2014). For logistic reasons, horseradish leaf samples were transported (1–2 h) in plastic bags to the lab as fresh and extracted directly, hence levels are reported on a fresh wt. basis. Extraction of plant material in boiling MeOH–H_2O (7:3) for 3 × 1 min, binding of GSLs to anion exchange mini columns and on-column enzymatic conversion to dGSLs as well as HPLC-UV followed Olsen et al. (2016). Results reported here represent analyses carried out between 2012 and 2018. Naturally, retention times varied over the years due to column history, as well as the change of column and gradient described below. However, each batch or group of consecutive batches included authentic references or in-house standard mixtures allowing the reported degree of either conclusive or tentative peak identification. Ion trap HPLC-MS was carried out with an Agilent 1100 Series HPLC (Agilent Technologies, Germany) coupled to a Bruker HCT-Ultra ion trap mass spectrometer (Bruker Daltonics, Bremen, Germany) fitted with a Zorbax SB-C18 column (Agilent; 1.8 μm, 2.1 × 50 mm). The oven temperature was maintained at 35 °C. The mobile phases were: A, water with 0.1% (v/v) HCOOH and 10 μM NaCl; B, acetonitrile with 0.1% (v/v) HCOOH. The gradient program was: 0–1 min, isocratic 2% B; 1–8 min, linear gradient 2%–40% B; 8–9 min, linear gradient 40%–90% B; 9–11.5 min isocratic 90% B; 11.5–11.6 min, linear gradient from 90% B to 2% B; 11.60–14 min, isocratic 2% B. The flow rate was 0.2 mL/min but increased to 0.3 mL/min in the interval 11.1–13.5 min. The mass spectrometer was run with positive electrospray ionization, in Ultra Scan mode with a mass range of *m/z* 100–800. The capillary voltage was −4000 V and the drying gas (N_2) was set to 365 °C with a flow of 10 L/min. MS2 spectra were acquired in automatic mode using SmartFrag with the instrument default settings (Isolation width 4 *m/z*, MS/MS fragmentation amplitude 1V, start amplitude 30%, end amplitude 200%, acquisition time 40 ms). The eluent conditions, with a trace of added NaCl, were used consistently from instrument acquisition in 2006 to end of use in 2018, for a heavily used instrument dedicated to analysis of a wide variety of mainly glycosidic metabolites. No problems were ever observed that could be connected with this use of a trace of salt in the eluents, and no special precautions (such as flushing pump parts) were ever used for preventing such problems. The final decision to end operation was due to a need for expensive service of the ion trap MS and was not related to mechanical LC parts, and the described routine use of 10 μM NaCl dope is recommended for dedicated instruments.

Chromatograms were manually investigated, including MS/MS analysis of each visually detected peak, systematic searches for *m/z* values of candidate dGSLs (as Na^+ adducts), and scanning the MS2 trace for characteristic fragments from acylated GSLs (e.g. *m/z* 323 for benzoylated and *m/z* 395 for isoferuloylated, in both cases representing the Na^+ adduct of acylated anhydroGlc). Statements of GSLs searched for but not detected were based on careful manual inspection of numerous samples representing all sample types/plant parts relevant for the case, supported by knowledge of the t_R of authentic standards (Table 2).

Major peaks were quantified from UV peak areas at 229 nm, using established relative response factors (Buchner, 1987), compared to the area of an external std. of potassium 2-propenylGSL treated like the samples. Minor peaks only detected in HPLC-MS (typically methylsulfinylalkyls and methylthioalkyls) were quantified from the signal intensity of the peak to be quantified divided by the signal intensity of a moderate intensity peak of comparable chemistry (typically Buen or an ω-(methylsulfinyl)alkylGSL). This relative MS signal intensity was multiplied with the level of the latter peak from UV to give the estimated level of the peak to be quantified. Controls using both UV and MS quantification of 6mSOhex showed reasonable correspondence of UV and MS based levels (slope 0.6 or 1.2, $R^2 = 0.8$ or 0.9 with and without an apparent outlier). IM was an exception, with a slope of 2.24 ($R^2 = 0.99$, $N = 14$). Hence, for trace indole GSLs estimated by MS, the level calculated from the peak intensity was divided by 2.24 to correct for apparently higher ionization efficiency of the indole GSLs.

4.4. Glucosinolate analysis by UHPLC-QToF of desulfoglucosinolates

Desulfated glucosinolate preparations as described in Section 4.3. were in a few cases also analyzed by UHPLC-QToF MS/MS analysis (UltraHigh Performance Liquid Chromatography Quadropole Time-of-Flight tandem mass spectrometry). This was the case for authentic desulfo 3mSOp for confirmation of the identity of the major peak in *I. amara* seed extracts, for confirmation of 4hEBAR, 4mSOb and 5mSOp and distinction of authentic 2h3mPe from an unknown isomer in *P. virginica* (Supplementary Fig. S4), for isomer distinction of isomeric hexyl and heptyl GSLs in *N. officinale* (Supplementary Fig. S5) and for characterization of an unidentified hydroxybutylGSL in *R. luteola* (Section 2.4.). UHPLC-QToF was performed using a Dionex Ultimate 3000RS UHPLC system (Thermo Fisher Scientific) system coupled to a

Compact™ (QToF) mass spectrometer (Bruker Daltonics) with an electrospray ionization source. The UHPLC system was equipped with a DAD detector, temperature controlled auto-sampler (10 °C) and column oven (set to 40 °C).

Either 5 or 10 μL sample was injected onto a Kinetex XB-C18 UHPLC column (100 × 2.1 mm, 1.7 μm, 100 Å pore size; Phenomenex) and eluted with a flow rate of 0.3 mL/min. Neither t_R nor MS was influenced by injection volume in this range (results not shown), allowing a moderate increase of sensitivity by using 10 μL injection of rather dilute samples or for repeated analysis with focus on minor peaks. The mobile phase comprised of a gradient of solvent A (0.05% formic acid in water) and solvent B (0.05% formic acid in acetonitrile). The initial composition was 98% A and 2% B, which was held for 2 min, before a linear increase to 40% solvent B over 18 min. Solvent B composition was increased to 90% B over a further 5 min and held at 90% solvent B for 5 min. The solvent composition was rapidly decreased back to 2% B over 1 min, and the column re-equilibrated at 2% B for 6 min prior to the next injection, giving a total method run time of 37 min.

The QToF mass spectrometer was operated in full scan positive ion mode with the following instrument settings: *m/z* 50–900: nebulizer gas (N_2), 2.0 bar; drying gas (N_2), 8 L/min; drying gas temperature, 220 °C; capillary voltage, 4000 V; spectra acquisition rate, 6 Hz. Data dependent MS/MS acquisition (collision energy was varied between 18 and 45 eV with increasing *m/z* of the precursor ion) was triggered for the three most intense ions in the *m/z* 130–900 region of the full scan MS spectra. Every chromatogram was calibrated to give accurate mass by automated infusion of sodium formate clusters (10 mM) at the beginning of each run.

All data acquisition was automated using a combination of Chromeleon Express (Thermo Fisher Scientific), Compass oTOF Control (Version 4.0.15.3248, Bruker Daltonics) and Hystar (Version 3.2 SR4, Bruker Daltonics) software. For data analysis, Compass DataAnalysis software (Version 4.3, Bruker Daltonics) was used. For measurement of high-resolution masses, we averaged detected values from available peak maxima.

4.5. DNA sequence analysis and phylogenetic reconstructions

Total DNA was extracted from 100 mg of dried leaf tissue from six *Barbarea vulgaris* accessions that were well-characterized for chemotype (P- and G-type) (B44, B59, B4, B54, B55, B56, Supplementary Table S1) using the Invisorb Spin Plant Mini Kit (Stratec Biomedical AG, Birkenfeld, Germany). PCR amplification of the internal transcribed spacer regions 1 and 2 of nuclear encoded ribosomal DNA (ITS) and including the 5.8 S rDNA gene was performed in a volume of 25 μL, using 10 μM of each primer, respectively, a total of 2.0 mM $MgCl_2$ and 0.5 U of Mango-Taq polymerase (Bioline, Luckenwalde, Germany). The primers used for amplification were *ITS*-for: 5′- GGAAGGAGAAGTCGTAACAA-3′, and *ITS*-rev: 5′- TCCTCCGCTTATTGATATGC-3′). All primers were extended by the M13 sequence for subsequent sequencing using M13 universal sequencing primers. The amplifications were run on a PTC 200 Peltier Thermal Cycler (MJ Research, Waltham, Massachusetts, USA) under the following conditions: 3 min initial denaturation at 95 °C; 30 cycles of amplification with 30 s at 95 °C, 30 s at 50 °C, and 1 min at 72 °C; and 5 min of final elongation at 72 °C. PCR success was checked with electrophoresis in a 1% agarose gel in TAE-buffer. PCR product clean-up was executed using the Wizard SV Gel and PCR Clean-Up System (Promega, Madison, USA). Custom Sanger-sequencing was performed with GATC-Biotech (Konstanz, Germany). The electropherograms were checked and trimmed to the borders of the analyzed markers using the program SeqMan DNA-Star Lasergene software package (DNASTAR, Madison, Wisconsin, USA).

Detailed insights into phylogenetic relationships within genus *Barbarea* were analyzed using SplitsTree version 4.8 (Huson and Bryant, 2006), based on aligned ITS sequences (comprising the internal transcribed spacer 1 and 2 of nuclear encoded ribosomal RNA and intervening 5.8 S rDNA gene). Analyses were run using the ParsimonySplits option and 1000 bootstrap replicates were calculated to infer further node support. The corresponding alignment is available as Supplementary Table S2. SplitsTree analysis was conducted with 75 ITS sequences from nine different *Barbarea* taxa. Detailed information on the various accessions is provided in Supplementary Table S1. The respective alignment is congruent with the enlarged tribal-wide (Cardamineae) analysis comprising 171 taxa (Agerbirk et al., 2021) and highlighting the phylogenetic position of *Barbarea*. This enlarged dataset comprises single representatives from *Barbarea* ITS types identified herein.

Author contributions

Conceived study: NA. Collected and identified plants: DC, MØ, TPH, NA. Cultivated and sampled plants: DC, NA. Performed insect-plant experiments and chemical induction: TPH, SC, KRJ. Analyzed GSLs: NA, CCH, CEO, DIP. ITS sequencing, phylogeny, split-tree analysis and related draft writing: MAK, CK. Contributed unpublished data related to ITS-sequences in GenBank and related discussion: CBAL. Wrote manuscript draft: NA. Discussed and edited manuscript: All.

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.

Acknowledgements

We thank Birgitte B. Rasmussen for skilled GSL extraction, desulfation and HPLC analysis, Dr. Renato Iori for a generous gift of 4mSbuen and 4mSObuen, Dr. Ulla L. Dolde for *A. thaliana* Col-0 seeds, Torben og Alice Frimodts Fond for financial support, and five anonymous reviewers for helpful and constructive comments to an earlier version of the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.phytochem.2021.112658.

References

Agerbirk, N., Ørgaard, M., Nielsen, J.K., 2003. Glucosinolates, flea beetle resistance, and leaf pubescence as taxonomic markers in the genus *Barbarea* (Brassicaceae). Phytochemistry 63, 69–80. https://doi.org/10.1016/S0031-9422(02)00750-1.

Agerbirk, N., Warwick, S., Hansen, P.R., Olsen, C.E., 2008. *Sinapis* phylogeny and evolution of glucosinolates and specific nitrile degrading enzymes. Phytochemistry 69, 2937–2949. https://doi.org/10.1016/j.phytochem.2008.08.014.

Agerbirk, N., Olsen, C.E., Chew, F., Ørgaard, M., 2010. Variable glucosinolate profiles of *Cardamine pratensis* (Brassicaceae) with equal chromosome numbers. J. Agric. Food Chem. 58, 4693–4700. https://doi.org/10.1021/jf904362m.

Agerbirk, N., Olsen, C.E., 2011. Isoferuloyl derivatives of five seed glucosinolates in the crucifer genus *Barbarea*. Phytochemistry 72, 610–623. https://doi.org/10.1016/j.phytochem.2011.01.034.

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., Olsen, C.E., Cipollini, D., Ørgaard, M., Linde-Laursen, I., Chew, F.C., 2014. Specific glucosinolate analysis reveals variable levels of glucobarbarins, dietary precursors of 5-phenyloxazolidine-2-thiones, in watercress types with contrasting chromosome number. J. Agric. Food Chem. 62, 9586–9596. https://doi.org/10.1021/jf5032795.

Agerbirk, N., Olsen, C.E., 2015. Glucosinolate hydrolysis products in the crucifer *Barbarea vulgaris* include a thiazolidine-2-one from a specific phenolic isomer as well as oxazolidine-2-thiones. Phytochemistry 115, 143–151. https://doi.org/10.1016/j.phytochem.2014.11.002.

Agerbirk, N., Olsen, C.E., Heimes, C., Christensen, S., Bak, S., Hauser, T., 2015. Multiple hydroxyphenethyl glucosinolate isomers and their tandem mass spectrometric distinction in a geographically structured polymorphism in the crucifer *Barbarea vulgaris*. Phytochemistry 115, 130–142. https://doi.org/10.1016/j.phytochem.2014.09.003.

Agerbirk, N., Matthes, A., Erthmann, P.Ø., Ugolini, L., Cinti, S., Lazaridi, E., Nuzillard, J.-M., Müller, C., Bak, S., Rollin, P., Lazzeri, 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., Kiefer, C., Hauser, T.P., Ørgaard, M., Lange, C.B.A., Cipollini, D., Koch, M.A., 2021. Comparison of glucosinolate diversity in the crucifer tribe Cardamineae and the remaining order Brassicales highlights repetitive evolutionary loss and gain of biosynthetic steps (in press) Phytochemistry (in press).

Agneta, R., Lelario, F., De Maria, S., Möllers, C., Bufo, S.A., Rivelli, A.R., 2014. Glucosinolate profile and distribution among plant tissues and phenological stages of field-grown horseradish. Phytochemistry 106, 178–187. https://doi.org/10.1016/j.phytochem.2014.06.019.

Badenes-Pérez, F.R., Parrado Marquez, B., Petitpierre, E., 2017. Can flowering *Barbarea* spp. Brassicaceae) be used simultaneously as a trap crop and in conservation biological control? J. Pest. Sci. 90, 623–633. https://doi.org/10.1007/s10340-016-0815-y.

Badenes-Pérez, F.R., Lopez-Perez, J.A., 2018. Resistance and susceptibility to powdery mildew, root-knot nematode, and Western flower thrips in two types of winter cress (Brassicaceae). Crop Protect. 110, 41–47. https://doi.org/10.1016/j.cropro.2018.03.015.

Bakhtiari, M., Glauser, G., Rasmann, S., 2018. Root JA induction modifies glucosinolate profiles and increases subsequent aboveground resistance to herbivore attack in *Cardamine hirsuta*. Front. Plant Sci. 9, 1230. https://doi.org/10.3389/fpls.2018.01230.

Bennett, R.N., Mellon, F.A., Kroon, P.A., 2004. Screening crucifer seeds as sources of specific intact glucosinolates using ion-pair high-performance liquid chromatography negative ion electrospray mass spectrometry. J. Agric. Food Chem. 52, 428–438. https://doi.org/10.1021/jf030530p.

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.

Brown, P.D., Tokuhisa, J.G., Reichelt, M., Gershenzon, J., 2003. Variation of glucosinolate accumulation among different organs and developmental stages of *Arabidopsis thaliana*. Phytochemistry 62, 471–481. https://doi.org/10.1016/s0031-9422(02)00549-6.

Byrne, S., Erthmann, P.Ø., Agerbirk, N., Bak, S., Hauser, T.P., Nagy, I., Paina, C., Asp, T., 2017. The genome sequence of *Barbarea vulgaris* facilitates the study of ecological biochemistry. Sci. Rep. 7, 40728. https://doi.org/10.1038/srep40728.

Cacho, N.I., Kliebenstein, D.J., Strauss, S.Y., 2015. Macroevolutionary patterns of glucosinolate defense and tests of defense-escalation and resource availability hypotheses. New Phytol. 208, 915–927. https://doi.org/10.1111/nph.13561.

Christensen, S., Heimes, C., Agerbirk, N., Kuzina, V., Olsen, C.E., Hauser, T.P., 2014. Different geographical distributions of two genotypes of *Barbarea vulgaris* that differ in resistance to insects and a pathogen. J. Chem. Ecol. 40, 491–501. https://doi.org/10.1007/s10886-014-0430-4.

Christensen, S., Sørensen, H., Munk, K.R., Hauser, T.P., 2016. A hybridisation barrier between two evolutionary lineages of Barbarea vulgaris (Brassicaceae) that differ in biotic resistances. Evol. Ecol. 30, 887–904. https://doi.org/10.1007/s10682-016-9858-z.

Christensen, S., Enge, S., Jensen, K.R., Müller, C., Kiær, L.P., Agerbirk, N., Heimes, C., Hauser, T.P., 2019. Different herbivore responses to two co-occurring chemotypes of the wild crucifer *B. vulgaris*. Arthropod Plant Inte 13, 19–30. https://doi.org/10.1007/s11829-018-9633-x.

Cole, R.A., 1976. Isothiocyanates, nitriles and thiocyanates as products of autolysis of glucosinolates in *Cruciferae*. Phytochemistry 15, 759–762. https://doi.org/10.1016/S0031-9422(00)94437-6.

Czerniawski, P., Piasecka, A., Bednarek, P., 2021. Evolutionary changes in the glucosinolate biosynthetic capacity in species representing *Capsella, Camelina* and *Neslia* genera. Phytochemistry 181, 112571. https://doi.org/10.1016/j.phytochem.2020.112571.

Daxenbichler, M.E., Spencer, G.F., Carlson, D.G., Rose, G.B., Brinker, A.M., Powell, R.G., 1991. Glucosinolate composition of seeds from 297 species of wild plants. Phytochemistry 30, 2623–2638. https://doi.org/10.1016/0031-9422(91)85112-D.

Dekic, M.S., Radulovic, N.S., Stojanovic, N.M., Randjelovic, P.J., Stojanovic-Radic, Z., Najman, S., Stojanovic, S., 2017. Spasmolytic, antimicrobial and cytotoxic activities of 5-phenylpentyl isothiocyanate, a new glucosinolate autolysis product from horseradish (*Armoracia rusticana* P. Gaertn., B. Mey. & Scherb., Brassicaceae). Food Chem. 232, 329–339. https://doi.org/10.1016/j.foodchem.2017.03.150.

Edger, P.P., Heidel-Fischer, H.M., Bekaert, M., Rota, J., Gloeckner, G., Platts, A.E., Heckel, D., Der, J.P., Wafula, E.K., Tang, M., Hofberger, J.A., Smithson, A., Hall, J.C., Blanchette, M., Bureau, T.E., Wright, S.I., dePamphilis, C.W., Schranz, M.E., Barker, M.S., Conant, G.C., Wahlberg, N., Vogel, H., Pires, J.C., Wheat, C.W., 2015. The butterfly plant arms-race escalated by gene and genome duplications. Proc. Natl. Acad. Sci. U.S.A. 112, 8362–8366. https://doi.org/10.1073/pnas.1503926112.

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., Sytsma, 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.

Erthmann, P.Ø., Agerbirk, N., Bak, S., 2019. A tandem array of UDP-glucosyltransferases from the UGT73C subfamily glycosylate sapogenins, forming a spectrum of Mon- and bisdesmosidic saponins. Plant Mol. Biol. 97, 37–55. https://doi.org/10.1007/s11103-018-0723-z.

Fahey, J.W., Zalcmann, A.T., Talalay, P., 2001. The chemical diversity and distribution of glucosinolates and isothiocyanates among plants. Phytochemistry 56, 5–51. https://doi.org/10.1016/s0031-9422(00)00316-2.

Gmelin, R., Kjær, A., Schuster, A., 1970. Glucosinolates in seeds of *Sibara virginica* (L.) Rollins: two new glucosinolates. Acta Chem. Scand. 24, 3031–3037. https://doi.org/10.3891/acta.chem.scand.24-3031.

Graeff, M., Straub, D., Eguen, T., Dolde, U., Rodrigues, V., Brandt, R., Wenkel, S., 2016. Microprotein-mediated recruitment of CONSTANS into a TOPLESS trimeric complex represses flowering in Arabidopsis. PLoS Genet. 12, e1005959 https://doi.org/10.1371/journal.pgen.1005959.

Griffiths, D.W., Deighton, N., Birch, A.N.E., Patrian, B., Baur, R., Städler, E., 2001. Identification of glucosinolates on the leaf surface of plants from the Cruciferae and other closely related plants. Phytochemistry 57, 693–700. https://doi.org/10.1016/s0031-9422(01)00138-8.

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.

Hartvig, P., 2015. Atlas Flora Danica. Gyldendal, Copenhagen.

Hegi, G., 1958. Illustrierte Flora von Mittel-Europa. Carl Hanser Verlag, Münich, Germany.

Heimes, C., Agerbirk, N., Sørensen, H., van Mölken, T., Hauser, T.P., 2016. Ecotopic differentiation of two sympatric chemotypes of *Barbarea vulgaris* (Brassicaceae) with different biotic resistances. Plant Ecol. 217, 1055–1068. https://doi.org/10.1007/s11258-016-0631-8.

Hussain, M., Gao, J., Bano, S., Wang, L., Lin, Y., Arthurs, S., Qasim, M., Mao, R., 2020. Diamondback moth larvae trigger host plant volatiles that lure its adult females for oviposition. Insects 11, 725. https://doi.org/10.3390/insects11110725.

Huson, D.H., Bryant, D., 2006. Application of phylogenetic networks in evolutionary studies. Mol. Biol. Evol. 23, 254–267. https://doi.org/10.1093/molbev/msj030.

Iori, R., Barillari, J., Gallienne, E., Bilardo, C., Tatibouët, A., Rollin, P., 2008. Thio-functionalised glucosinolates: unexpected transformation of desulfoglucoraphenin. Tetrahedron Lett. 49, 292–295. https://doi.org/10.1016/j.tetlet.2007.11.059.

Jeon, J., Bong, S.J., Park, J.S., Park, Y.K., Arasu, M.V., Al-Dhabi, N.A., Park, S.U., 2017. De novo transcriptome analysis and glucosinolate profiling in watercress (*Nasturtium officinale* R. Br.). BMC Genom. 18, 401. https://doi.org/10.1186/s12864-017-3792-5.

Kakizaki, T., Kitashiba, H., Zou, Z., Li, F., Fukino, N., Ohara, T., Nishio, T., Ishida, M., 2017. A 2-oxoglutarate-dependent dioxygenase mediates the biosynthesis of glucoraphasatin in radish. Plant Physiol. 173, 1583–1593. https://doi.org/10.1104/pp.16.01814.

Kjær, A., 1976. Glucosinolates in the cruciferae. In: Vaughan, J.G., MacLeod, A.J., Jones, B.M.G. (Eds.), The Biology and Chemistry of the Cruciferae. Academic Press, London, pp. 207–219.

Kjær, A., Gmelin, R., 1958. Isothiocyanates XXXIII. An isothiocyanate glucoside (glucobarbarin) of *Reseda lueteola* L. Acta Chem. Scand. 12, 1693–1694.

Kliebenstein, D.J., Kroyman, J., Brown, P., Figuth, A., Pedersen, D., Gershenzon, J., Mitchell-Olds, T., 2001. Genetic control of natural variation in *Arabidopsis* glucosinolate accumulation. Plant Physiol. 126, 811–825. https://doi.org/10.1104/pp.126.2.811.

Knill, T., Schuster, J., Reichelt, M., Gershenzon, J., Binder, S., 2008. Arabidopsis branched-chain aminotransferase 3 fuctions in both amino acid and glucosinolate biosynthesis. Plant Physiol. 146, 1028–1039. https://doi.org/10.1104/pp.107.111609.

Kumar, R., Lee, S.G., Augustine, R., Reichelt, M., Vassão, D.G., Palavalli, M.H., Allen, A., Gershenzon, J., Jez, J.M., Bisht, N.C., 2019. Molecular basis of the evolution of methylthioalkylmalate synthase and the diversity of methionine-derived glucosinolates. Plant Cell 31, 1633–1647. https://doi.org/10.1105/tpc.19.00046.

Lange, C.B.A., Hauser, T.P., Deichmann, V., Ørgaard, M., in review. Hybridization and complex evolution of *Barbarea vulgaris* and related species (Brassicaceae). Mol. Phylogenet. Evol.

Larsen, L.M., Olsen, O., Sørensen, H., 1983. Failure to detect glucosinolates in *Plantago* species. Phytochemistry 22, 2314–2315. https://doi.org/10.1016/S0031-9422(00)80170-3.

Liu, T., Zhang, X., Yang, H., Agerbirk, N., Qiu, Y., Wang, H., Shen, D., Song, J., Li, X., 2016. Aromatic glucosinolate biosynthesis pathway in *Barbarea vulgaris* and its response to *Plutella xylostella* infestation. Front. Plant Sci. 7, 83. https://doi.org/10.3389/fpls.2016.00083.

Liu, T.-J., Zhang, Y.-J., Agerbirk, N., Wang, H.-P., Wei, X.-C., Song, J.-P., He, H.-J., Zhao, X.-Z., Zhang, X.-H., Li, X.-X., 2019a. A high-density genetic map and QTL mapping of leaf traits and glucosinolates in *Barbarea vulgaris*. BMC Genom. 20, 371. https://doi.org/10.1186/s12864-019-5769-z.

Liu, Q., Khakimov, B., Cardenas, P.D., Cozzi, F., Olsen, C.E., Jensen, K.R., Hauser, T.P., Bak, S., 2019b. The cytochrome P450 CYP72A552 is key to production of hederagenin-based saponins that mediate plant defense against herbivores. New Phytol. 222, 1599–1609. https://doi.org/10.1111/nph.15689.

Liu, X., Zhang, L., Feng, X., Lv, B., Li, C., 2017. Biosynthesis of glycyrrhetinic acid-3-*O*-monoglucose using glycosyltransferase UGT73C11 from *Barbarea vulgaris*. Ind. Eng. Chem. Res. 56, 14949–14958. https://doi.org/10.1021/acs.iecr.7b03391.

Mandákova, T., Lysak, M.A., 2019. Healthy roots and leaves: comparative genome structure of horseradish and watercress. Plant Physiol. 179, 66–73. https://doi.org/10.1104/pp.18.01165.

Melichárková, A., Šlenker, M., Zozomová-Lihová, J., Skokanová, K., Šingliarová, B., Kačmárová, T., Caboňová, M., Kempa, M., Šrámková, G., Mandáková, T., Lysák, M. A., Svitok, M., Mártonfiová, L., Marhold, K., 2020. So closely related and yet so different: strong contrasts between the evolutionary histories of species of the *Cardamine pratensis* popyploid species complex in Central Europe. Front. Plant Sci. 11, 588856. https://doi.org/10.3389/fpls.2020.588856.

Mithen, R., Bennett, R., Marquez, J., 2010. Glucosinolate biochemical diversity and innovation in the Brassicales. Phytochemistry 71, 2074–2086. https://doi.org/10.1016/j.phytochem.2010.09.017.

Montaut, S., Bleeker, R.S., 2013. Review of *Cardamine diphylla* (michx.) A. Wood (Brassicaceae): ethnobotany and glucosinolate chemistry. J. Ethnopharmacol. 149, 401–408. https://doi.org/10.1016/j.jep.2013.07.020.

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 laciniata* all.: unveiling 4'-*O*-*β*-D-apiofuranosylglucomatrolin. Carbohydr. Res. 488, 107898. https://doi.org/10.1016/j.carres.2019.107898.

Olsen, C.E., Huang, X.-C., Hansen, C.I.C., Cipollini, D., Ørgaard, M., Matthes, A., Geu-Flores, F., Koch, M.A., Agerbirk, N., 2016. Glucosinolate diversity within a phylogenetic framework of the tribe Cardamineae (Brassicaceae) unraveled with HPLC-MS/MS and NMR-based analytical distinction of 70 desulfoglucosinolates. Phytochemistry 132, 33–56. https://doi.org/10.1016/j.phytochem.2016.09.013.

Pagnotta, E., Agerbirk, N., Olsen, C.E., Ugolini, L., Cinti, S., Lazzeri, L., 2017. Hydroxyl and methoxyl derivatives of benzylglucosinolate in *Lepidium densiflorum* with hydrolysis to isothiocyanates and non-isothiocyanate products: substitution governs product type and mass spectral fragmentation. J. Agric. Food Chem. 65, 3167–3178. https://doi.org/10.1021/acs.jafc.7b00529.

Pagnotta, E., Montaut, S., Matteo, R., Rollin, P., Nuzillard, J.M., Lazzeri, 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.

Pedras, M.S.C., Alavi, M., To, Q.H., 2015. Expanding the nasturlexin family: nasturlexins C and D and their sulfoxides are phytoalexins of the crucifers *Barbarea vulgaris* and *B. verna*. Phytochemistry 118, 131–138. https://doi.org/10.1016/j.phytochem.2015.08.009.

Pedras, M.S.C., To, Q.H., 2018. Interrogation of biosynthetic pathways of the cruciferous phytoalexins nasturlexins with isotopically labelled compounds. Org. Biomol. Chem. 16, 3625–3638. https://doi.org/10.1039/C8OB00673C.

Petersen, A., Hansen, L.G., Mirza, N., Crocoll, C., Mirza, O., Halkier, B.A., 2019. Changing substrate specificity and iteration of amino acid chain elongation in glucosinolate biosynthesis through targeted mutagenesis of *Arabidopsis* methylthioalkylmalate synthase 1. Biosci. Rep. 39, BSR20190446 https://doi.org/10.1042/BSR20190446.

Pfalz, M., Mukhaimar, M., Perreau, F., Kirk, J., Hansen, C.I.C., Olsen, C.E., Agerbirk, N., Kroymann, J., 2016. Methyl transfer in glucosinolate biosynthesis mediated by indole glucosinolate *O*-methyltransferase 5. Plant Physiol. (Wash. D C) 172, 2190–2203. https://doi.org/10.1104/pp.16.01402.

Rapo, C.B., Schaffner, U., Eigenbrode, S.D., Hinz, H.L., Price, W.J., Morra, M., Gaskin, J., Schwarzlände, M., 2019. Feeding intensity of insect herbivores is associated more closely with key metabolite profiles than phylogenetic relatedness of their potential hosts. PeerJ 7, e8203. https://doi.org/10.7717/peerj.8203.

Reichelt, M., Brown, P.D., Schneider, B., Oldham, N.J., Stauber, E., Tokuhisa, J., Kliebenstein, D.J., Mitchell-Olds, T., Gershenzon, J., 2002. Benzoic acid glucosinolate esters and other glucosinolates from *Arabidopsis thaliana*. Phytochemistry 59, 663–671. https://doi.org/10.1016/s0031-9422(02)00014-6.

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.

Sun, R., Jiang, X., Reichelt, M., Gershenzon, J., Pandit, S.S., Vassão, D.G., 2019. Tritrophic metabolism of plant chemical defenses and its effect on herbivore and predator performance. eLife 8, e51029. https://doi.org/10.7554/eLife.51029.

Textor, S., de Kraker, J.-W., Hause, B., Gershenzon, J., Tokuhisa, J.G., 2007. MAM3 catalyzes the formation of all aliphatic glucosinolate chain lengths in Arabidopsis. Plant Physiol. 144, 60–71. https://doi.org/10.1104/pp.106.091579.

Toneatto, F., Nielsen, J.K., Ørgaard, M., Hauser, T.P., 2010. Genetic and sexual separation between insect resistant and susceptible *Barbarea vulgaris* plants in Denmark. Mol. Ecol. 19, 3456–3465. https://doi.org/10.1111/j.1365-294X.2010.04760.x.

Toneatto, F., Hauser, T.P., Nielsen, J.K., Ørgaard, M., 2012. Genetic diversity and similarity in the *Barbarea vulgaris* complex (Brassicaceae). Nord. J. Bot., Le 30, 506–512. https://doi.org/10.1111/j.1756-1051.2012.01546.x.

Visentin, M., Tava, A., Iori, R., Palmieri, S., 1992. Isolation and identification of *trans*-4-(methylthio)-3-butenyl glucosinolate from radish roots (*Raphanus sativus* L.). J. Agric. Food Chem. 40, 1687–1691. https://doi.org/10.1021/jf00021a041.

Voutsina, N., Payne, A.C., Hancock, R.D., Clarkson, G.J.J., Rothwell, S.D., Chapham, M. A., Taylor, G., 2016. Characterization of the watercress (*Nasturtium officinale* R. Br.; Brassicaceae) transcriptome using RNASeq and identification of candidate genes for important phytonutrient traits linked to human health. BMC Genom. 17, 378. https://doi.org/10.1186/s12864-016-2704-4.

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. https://doi.org/10.1101/2020.05.12.090720 (in press).

Windsor, A.J., Reichelt, M., Figuth, A., Svatos, A., Kroymann, J., Kliebenstein, D.J., Gershenzon, J., Mitchell-Olds, T., 2005. Geographic and evolutionary diversification of glucosinolates among near relatives of *Arabidopsis thaliana* (Brassicaceae). Phytochemistry 66, 1321–1333. https://doi.org/10.1016/j.phytochem.2005.04.016.

Zidorn, C., 2017. Guidelines for consistent characterisation and documentation of plant source materials for studies in phytochemistry and phytopharmacology. Phytochemistry 139, 56–59. https://doi.org/10.1016/j.phytochem.2017.04.004.

J.-M. Züst, T., Strickler, S.R., Powell, A.F., Mabry, M.E., An, H., Mirzaei, M., York, T., Holland, C.K., Kumar, P., Erb, M., Petschenka, G., Gómez, J.-M., Perfectti, F., Müller, C., Pires, J.C., Mueller, L.A., Jander, G., 2020. Independent evolution of ancestral and novel defenses in a genus of toxic plants (*Erysimum*, Brassicaceae) eLife 9, e51712 https://doi.org/10.7554/eLife.51712.