JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY ARTICLE

DOI:10.1021/jf904362m

Variable Glucosinolate Profiles of *Cardamine pratensis* (Brassicaceae) with Equal Chromosome Numbers

NIELS AGERBIRK,*,[†] CARL ERIK OLSEN,[†] FRANCES S. CHEW,[§] AND MARIAN ØRGAARD[†]

[†]Faculty of Life Sciences, University of Copenhagen, Thorvaldsensvej 40, 1871 Frederiksberg, Denmark, and [§]Department of Biology, Tufts University, Medford, Massachusetts 02155

A novel glucosinolate, 3-(hydroxymethyl)pentylglucosinolate, was present at high levels in *Cardamine pratensis* L. from eastern North America and in commercially obtained seeds, but not in *C. pratensis* plants from southern Scandinavia. Glucosinolates in a number of accessions of *C. pratensis* included glucosinolates with the side chains 1-methylethyl, 1-(hydroxymethyl)ethyl, 1-methylpropyl, 1-(hydroxymethyl)propyl, 3-methylpentyl, 3-(hydroxymethyl)pentyl, benzyl, 4-hydroxybenzyl, 4-methoxybenzyl, indol-3-ylmethyl (as well as its 1-methoxy, 4-hydroxy, and 4-methoxy derivatives) and the rare side chain 1,4-dimethoxyindol-3-ylmethyl. Substantial variation was observed for four biosynthetic characters: (i) extent of chain elongation of Ile-derived glucosinolates; (ii) biosynthesis of Phe/Tyr-derived glucosinolates in general; (iii) hydroxylation of branched-chain glucosinolates; and (iv) *O*-methylation of 4-hydroxybenzylglucosinolate (sinalbin). Cytological analysis of pollen mother cells and root tip cells in meiosis and mitosis established the chromosome number to be $2n = 30$ for all accessions, irrespective of glucosinolate profile.

KEYWORDS: Glucosinolate polymorphism; side-chain hydroxylation; *O*-methylation; 3-(hydroxymethyl)-pentylglucosinolate; 4-hydroxybenzylglucosinolate; ploidy level; NMR spectra

INTRODUCTION

Glucosinolates are amino acid-derived secondary metabolites that are part of an activated defense system in plants of the mustard-cabbage order (Brassicales) (*1*). Upon tissue disruption, the glucosinolates are hydrolyzed by endogenous thioglucosidase enzymes to defensive products. Depending on the precursor amino acid and subsequent modifications, as well as a number of additional specifier proteins, the chemical and biological properties of the degradation products vary tremendously. For this reason, the continued search for additional structures and bioactivities has been a priority in this field for decades (see, e.g. refs *2*−*5*).

The most recent list of known glucosinolate structures in a critical review provided a stimulating overview of known glucosinolate biodiversity, and the authors suggested continued investigations of wild species (*6*). Further investigations have resulted in several new, interesting structures: some novel compounds were revealed after extensive taxonomic sampling in less well studied taxonomic groups (*7*), whereas other novel structures were revealed after detailed investigation of one or a few species (*8*−*10*). A notable example of the latter approach is the impressive list of known as well as unexpected structures revealed by the highly focused investigations of the model plant *Arabidopsis thaliana* (*11*). Detailed investigation of a single species has the advantage that intraspecific variation may be revealed, a feature that opens possibilities for genetic and ecological studies. Examples of highly variable species that have allowed genetic and ecological investigations of glucosinolate effects include *Arabidopsis thaliana* (*12*), *Barbarea vulgaris* (*13*, *14*), *Boechera stricta* (*15*), and *Brassica oleracea* (*16*).

Whereas detailed investigation of a single species usually involves complete structural elucidation, screening programs may involve a more superficial identification, frequently by chromatography coupled to MS (*2*, *7*). For example, a novel phenolic glucosinolate, homosinalbin, was predicted from a screening based on HPLC-MS (*7*) and later confirmed spectroscopically (*17*). In the same screening a number of unusual structures and unidentified peaks were suggested from seeds of the wild plant *Cardamine pratensis* (*7*). In connection with an investigation of native and introduced North American host plants of *Pieris* butterflies (*18*), we screened glucosinolate profiles of several species and detected a dominant peak with unexpected molecular mass in *C. pratensis*. As *C. pratensis* is a variable species with extremely complex phylogeny, recent geographic expansion, and a number of known ploidy levels, we decided to investigate the potential glucosinolate variation within this species compared to ploidy level and geographical origin.

We here report the structural elucidation of a novel glucosinolate from some accessions of this species and extensive polymorphisms in glucosinolate profiles within Scandinavian material and between these and eastern North American and other accessions. We further demonstrate by cytological analysis that the variation occurs among plants with the same chromosome number, $2n = 30$.

MATERIALS AND METHODS

Identification and Sampling of Plants. Flowering *C. pratensis* for sampling in the field was identified by N.A. (Danish plants) and F.S.C.

*Corresponding author (telephone +45 35 33 24 38; fax +45 35 33 23 98; e-mail nia@life.ku.dk).

© 2010 American Chemical Society
Published on Web 03/24/2010

(Massachusetts plants) from the characteristic morphology of flowering plants. One flowering U.S. (Massachusetts) specimen of *C. pratensis* was photographed and, on the basis of this photograph, the identity confirmed by I. Al-Shehbaz, Missouri Botanical Garden, St. Louis, MO. Additional Danish *C. pratensis* specimens from all three investigated localities were transplanted to pots in the fall of 2008 and vernalized outdoors, and the identity was confirmed by M.Ø. at flowering. Voucher specimens of the Danish plants are deposited at CP (the herbarium of the Faculty of Life Sciences, University of Copenhagen); voucher specimens of American plants will be deposited at CP after flowering in 2010.

Plants for chemical analysis were collected in plastic bags at field sites in 2008 (from flowering plants in May unless otherwise noted) and transported to the laboratory for dissection and lyophilization within 2−4 h, with the exception of *C. pratensis* collected at locality DK3 (plants a and b), in which case plant material was dried in the shade during a warm, sunny day and transported to the laboratory within 30 h. The localities were MA1, a maple−hemlock forest near the Housatonic River between Lee and New Lenox, MA (two collections: May 2005 and August 2007); MA2, a sandy, nonforested area that is seasonally flooded by the Housatonic River, 0.8 km straight-line-distance from MA1, near New Lenox, MA; DK1, mowed lawns at Lake Utterslev Mose, Copenhagen, Denmark (public lawns 0.2−0.8 km northeast of the lake and northwest of Highway 16); DK2, a dry, ungrazed meadow at Lake Gentofte Sø, Copenhagen, Denmark; and DK3, a water-logged soil, permanently or periodically submerged, at the rim of the minute pond "Mesing Gadekær", Skanderborg, Denmark (three individual plants, numbered DK3 a/b/c were sampled). Commercially obtained *C. pratensis* seeds (COM) were obtained from Chiltern Seeds, Ulverston, U.K., but upon inquiry, Jelitto Perennial Seeds Ltd., Schwarmstedt, Germany, was indicated as the source of the delivered seeds. The latter supplier indicated that the seeds were probably of southern German origin. Plants were grown from the seeds in an enriched peat-based substrate in a greenhouse with supplemental heat and light.

For chromosome counts, Danish rosette plants taken from the localities were cultivated outdoors in pots. Roots tips were taken from the bottom of the pots, and flower buds were collected in the spring for preservation in EtOH/AcOH. After sampling of flower buds, plants from most accessions (DK1/2/3 and COM) were cultivated outdoors on a bench, with labels indicating identity. When successful cytological investigation was accomplished, the exact individuals for which chromosome counts were obtained were sampled (rosette leaves) for glucosinolate analysis in September, after several months of cocultivation. In the case of plants from locality MA2, intact plants were collected in April 2009 as rootstock and kept in a cold room (ca. 5 °C) until mid June 2009, when they were brought out to room temperature and the plants flowered; buds of individual plants were collected and preserved in EtOH/AcOH, and upper leaves of the same individuals were lyophilized within 10 min and later analyzed for glucosinolates.

Determination and Identification of Glucosinolates. Glucosinolates were determined as desulfo derivatives (indicated by "d" before glucosinolate number) by HPLC with diode array detection (HPLC-DAD) (*17*). We used generally accepted relative response factors (rrfs)(*19*) or assumed rrfs based on side-chain structures (**9** was assumed to have the same rrf as **8**, **14** was assumed to have the same rrf as **11**, and aliphatic glucosinolates for which rrfs were not available (**1**−**6**, iso**6**) were assumed to have the rrfs of 1.0). In general, peak identities were based on HPLC-MS of at least one sample of each accession (including two samples with **14**) and by comparison with the authentic standards isolated as described below or reported previously (*8*, *17*). The column stationary phase and temperature were optimized due to coelution of d**3** and d**6** in *C. pratensis* samples at our usual conditions (*17*). Instead, we employed a 250 mm × 4.6 mm i.d., 5 μm, Luna phenyl hexyl column with a 3 mm Security Guard filter with a phenylpropyl stationary phase from Phenomenex (Torrance, CA), operated at 38 °C with a gradient program consisting of 2 min of water followed by a linear gradient of MeOH from 0 to 60% in 48 min, followed by a brief wash with MeOH and equilibration with water. Typical retention times (min) were as follows: d**2**, 6.9; d**4**, 13.5; d**1**, 14.6; d**8**, 15.6; d**6**, 20.0; d**12**, 20.4; d**3**, 21.4; d-iso**6**, 24.9; d**7**, 26.8; d**10**, 29.2; d**9**, 29.9; d**13**, 34.6; d**5**, 37.7; d**11**, 41.2; d**14**, 45.5. As there was an overlap of d**6** and d**12**, samples with d**6** were in addition analyzed for d**12** with the conditions described elsewhere (*17*): d**3** + d**6**, 18.8 min; d**12**, 21.2 min. Selected peaks

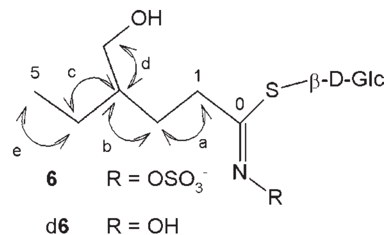

Figure 1. Novel glucosinolate, 3-(hydroxymethyl)pentylglucosinolate (**6**), from eastern North American *Cardamine pratensis*. Identification was based on the structure of the desulfated derivative (d**6**). Curved arrows indicate coupling revealed by COSY; letters of the various interactions refer to **Table 1**.

(d**1**, d**2**, d**3**, d**4**, d**6**, d**8**) were identified by NMR and HPLC-MS (*20*) after isolation by preparative HPLC from material from accessions MA1, DK1, and DK3a, and d**6**, d-iso**6**, and d**14** were characterized by high-resolution MS (*20*) of sodium adducts $[M + Na]^+$ of desulfo derivatives of glucosinolates from MA1 foliage and COM roots. In the case of DK1, used for preparation of the low-level peaks d**1**, d**2**, d**3**, d**4**, and dominating d**8**, extract of 1 g of dry weight of foliage collected in September was distributed over five analytical (0.1 g) DEAE columns and desulfo derivatives were prepared (*17*). The 25 mL combined eluent was lyophilized to dryness and the remnant dissolved in 1.1 mL of H_2O, which was subjected to HPLC with the usual analytical phenylhexyl column and gradient (see above). We injected 200 μL of concentrated sample per run, corresponding to at least 100 times more desulfoglucosinolate than a typical analytical injection, resulting in massive detector saturation for the dominating peak but still almost baseline separation of all peaks including d**4**, d**1**, and d**8** eluting during 4 min. A piece of short tubing had been attached at the DAD outlet, and individual peaks were collected manually, guided by the detector signal, taking a 5 s delay into account. The MeOH content of fractions was evaporated under an air stream, followed by freezing and lyophilization of the remaining fractions, which were dissolved in D_2O and subjected to NMR and HPLC-MS.

Chromosome Number Determination. Numbers of chromosomes were counted either in 4′,6-diamidino-2-phenylindole (DAPI)-stained pollen mother cells (PMCs) at meiotic stages from diakinesis to telophase II or in interphase nuclei and cells at somatic metaphase of root tips or tapetum cells of anthers. PMCs were derived from buds fixed in absolute EtOH/AcOH (3:1) for 12 h at room temperature and for a minimum of 12 h at −18 °C until use. After transfer to a drop of 45% aqueous AcOH on a chrome-acetic-cleaned slide, the stamens were dissected out of the buds and squashed under a coverslip. Preparation of root tips followed the method fo Ørgaard et al. (*21*). Slides were frozen in liquid nitrogen, coverslips lifted off using a razor blade, and slides dried and DAPI stained ($1\ \mu g\ \mu L^{-1}$). Slides were examined with a Zeiss Axioskop at 1000× magnification with appropriate filters for DAPI. From each accession approximately five slides were made, each slide providing around 20 counts.

RESULTS

Identification of Glucosinolates. We discovered a novel glucosinolate, 3-(hydroxymethyl)pentylglucosinolate (**6**, **Figure 1**), in foliage from *C. pratensis* plants collected in Massachusetts in 2005 (locality MA1). The glucosinolate was identified on the basis of the structural elucidation of the desulfo derivative d**6**, obtained in a commonly employed protocol for glucosinolate analysis (*19*), whereas the possible accumulation of desulfated glucosinolates as biosynthetic intermediates in the plant was not addressed in this investigation. Standard spectroscopic experiments served to elucidate the structure of the side chain (**Tables 1** and **2**), with the exception of the chirality. The structure of the remaining part of the molecule was assumed to be the well-known general glucosinolate structure based on the close similarity of proton and carbon signals to spectra of known compounds (**Tables 1** and **2**) and agreement between high-resolution MS and the elucidated

Table 1. ^{1}H NMR Spectroscopic Data for Desulfo Derivatives of Selected Aliphatic Glucosinolates Detected in *Cardamine pratensis*[a]

	d2	d4	d6
thioglucose meoity			
1′	5.01 (1H, d, J = 10)	5.02 (1H, d, J = 10)	4.99 (1H, d, J = 10)
2′−5′	3.4−3.6 (4H, m)	3.4−3.6 (4H, m)	3.4−3.6 (4H, m)
6′a	3.69 (1H)	3.68 (1H)	3.70 (1H, dd, J = 12 and 6)
6′b	3.89 (1H, dd, J = 12 and 2)	3.88 (1H, dd, J = 13 and 2)	3.88 (1H, dd, J = 12 and 2)
side chain			
1	2.94[ab] (1H, hex, J = 7)	2.78[ab] (1H, quint, J = 7)	2.63[a] (2H, dd, J = 8)
2	1.21[a] (3H, trip, J = 7)	1.63[ac] (2H, m)	1.65[ab] (2H, quart., $J \cong$ 6−8)
3		0.93[c] (3H, trip, J = 7)	1.51[bcd] (1H, hep, $J \cong$ 6)
4			1.37[ce] (2H, quint, J = 7)
5			0.89[e] (3H, trip, J = 7)
−CH$_2$OH	3.65[b] (1H) 3.83[b] (1H, dd, J = 11 and 7)	3.70[b] (1H) 3.81[b] (1H, dd, J = 11 and 7)	3.50[d] (2H)

[a] Chemical shifts (in D$_2$O) as δ relative to dioxane (δ_H = 3.75), J in Hz; signals without J and multiplicity indicated were extracted from 2D spectra (HSQC, HMBC, COSY). [a−e], side chain signals marked with the same letter were coupled, as evidenced by COSY.

Table 2. ^{13}C NMR Spectroscopic Data for Selected Desulfated Glucosinolates Detected in *Cardamine pratensis*[a]

	d2	d4	d5	d6
thioglucose meoity				
1′	82.2	82.6	82.3	82.3
2′	nd	nd	73.0	73.0
3′	nd	nd	78.0	78.0
4′	nd	nd	70.0	70.0
5′	nd	nd	81.0	81.0
6′	61.5	61.8	61.4[b]	61.4[b]
side chain				
0	157.5	156.4	158.1	158
1	40.0	47.3	30.5[b]	30.1[b]
2	16.7	24.2	34.7[b]	28.8[b]
3		11.4	34.4	41.3
4			29.2[b]	23.2[b]
5			11.4	11.0
Me			19.2	
−CH$_2$−OH	65.2	64.2		64.4[b]
HPLC-MS				
t_R (min)	2.1	2.6	8.0	3.9
m/z of [M + Na]$^+$)	320	334	346	362

[a] Chemical shifts (in D$_2$O) as δ relative to dioxane (δ_C = 67.4). Some signals were extracted from 2D spectra. nd, not determined because reported chemical shifts were extracted from 2D spectra. [b] Signal confirmed to be due to CH$_2$ by DEPT, which was carried out for d**5** and d**6** only.

structure of d**6** (found, m/z 362.1231; calcd for $C_{13}H_{25}O_7NSNa^+$, m/z 362.1244). The existence of the corresponding intact glucosinolate **6** in the plant tissue was inferred from release of d**6** after treatment of native metabolites, bound to an anion exchange column, with the enzyme sulfatase.

Previous investigations of *C. pratensis* were of considerable age (*6*) and could be considered preliminary, except for a recent investigation that was restricted to seeds (*7*). Hence, we investigated some additional accessions, selected according to availability, and identified an additional five aliphatic glucosinolates (**1**−**5**). The aliphatic glucosinolates **1**−**6** constituted three pairs of branched side chain glucosinolates with and without hydroxy groups on the side chain methyl group, apparently derived from the amino acids Val, Ile, and dihomoisoleucine (dihomo-Ile), respectively (**Figure 2**). The NMR spectra of the additional aliphatic structures also served as a useful reference for the chemical shifts of d**6**; the observed chemical shift differences were in accordance with the elucidated structures (**Tables 1** and **2**). The ^{1}H NMR spectra of d**1** and d**3** were also obtained but are not reported because they were identical to previously reported spectra of the same compounds from other sources (*17*). However, the identification of d**5** relied solely on comparison with authentic standards; previously unpublished data from ^{13}C NMR of d**5** isolated from *Erucastrum canariense* Webb & Bertel (*17*) are reported in **Table 2** merely for comparison with the data for d**6**.

A dominant aromatic glucosinolate from Danish accessions was identified as **8** (4-hydroxybenzylglucosinolate, commonly known as sinalbin) by comparison of retention time, m/z value, and ^{1}H NMR spectrum of the desulfated derivative to those of an authentic standard (*17*). An additional seven aromatic glucosinolates, **7** and **9**−**14** (**Figure 2**), were readily identified, on the basis of UV spectra, m/z values, and retention times identical to those of authentic standards (*8*, *17*), so NMR confirmation was considered to be superfluous in these cases. To confirm the identification of the rare indole glucosinolate **14**, the observed tiny peak of d**14** was isolated and its identity confirmed by the UV spectrum with a characteristic local maximum at 301 nm and expected exact molecular mass (found, m/z 451.1145; calcd for $C_{18}H_{24}O_8N_2SNa^+$, m/z 451.1146).

Whereas the remaining, unidentified peaks in most samples were quantitatively negligible, this was not the case for samples from the commercial accession COM. From COM roots, we observed an HPLC-DAD peak with a UV spectrum as expected for an aliphatic desulfoglucosinolate and with mass/charge ratio as would be expected for the sodium adduct of an isomer of d**6** (measured m/z 362.1235; calcd for $C_{13}H_{25}O_7NSNa^+$, m/z 362.1244). The corresponding intact glucosinolate is referred to as "iso**6**" in the following. The desulfo iso**6** would also have the same nominal mass as desulfo 4-methylthiobut-3-enylglucosinolate (calcd m/z 362.0703 for $C_{12}H_{21}O_6NS_2Na^+$), but this alternative was rejected due to the deviating exact molecular mass and ^{1}H NMR of a small preparation showing a complicated coupling pattern without a double bond. From COM seeds, we registered peaks with the nominal masses 378, 348, and 332, possibly corresponding to a hydroxylated derivative of **6** and lower homologues of **6** and **5** or isomers (**Figure 3**).

Variable Glucosinolate Profiles in *C. pratensis*. The novel, aliphatic **6** was also dominating in a second, independent collection (pooled leaves from several plants collected in 2007) of *C. pratensis* from Massachusetts (locality MA1), as well as in commercially obtained seeds (COM) and in roots and leaves of plants grown from the seeds. However, analysis of some Danish accessions revealed that glucosinolate profiles of *C. pratensis* from various localities varied, not only between continents but also between three Danish localities (**Figure 4**; **Table 3**).

The most obvious difference was that the collected Danish plants (DK1−3) were relatively low in aliphatic glucosinolates

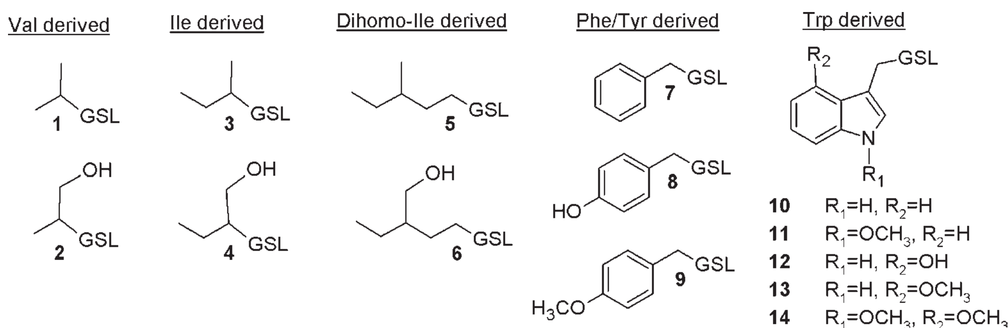

Figure 2. Glucosinolates detected in *Cardamine pratensis* from various localities in southern Scandinavia and eastern North America. GSL indicates the invariable part of the glucosinolate molecule, $C(SGlc)NOSO_3^-$. Dihomo-Ile = dihomoisoleucine. Other precursor abbreviations indicate standard amino acids.

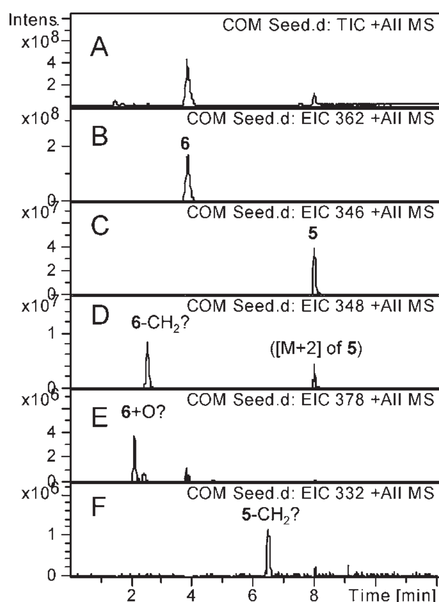

Figure 3. Demonstration of **5**, **6**, and three unidentified, putative glucosinolates by HPLC-MS analysis of desulfated derivatives from commercially obtained seeds of *C. pratensis*: (**A**) total ion chromatogram; (**B**−**F**) extracted ion chromatograms monitoring the *m/z* values 362 (**B**), 346 (**C**), 348 (**D**), 378 (**E**), and 332 (**F**). Sodium adducts of desulfo glucosinolate peaks are labeled with the numbers of the corresponding intact glucosinolates for simplicity. Peak designations followed by question marks refer to the discussion in the text and are entirely hypothetical.

but dominated by the Phe/Tyr derived **8** (or in one case **7** as described below), which was not detected in MA1 and COM plants. Two Danish accessions (DK1 nd DK2) also accumulated the *O*-methyl derivative **9**, which was likewise absent from MA1 and COM plants.

Danish plants were all devoid of **6** but did occasionally contain a trace of the structurally related **5**, which was also a minor glucosinolate in MA and COM plants. However, Danish plants contained additional aliphatic glucosinolates with shorter side chains apparently derived from Val and Ile. They were mainly oxidized (**2** and **4**) in plants from locality DK1 and mainly not oxidized (**1** and **3**) in plants from the nearby locality DK2, whereas plants from a third locality (DK3) were variable. Thus, Danish accessions showed no or only minor accumulation of dihomo-Ile-derived glucosinolates and varied between one another in the extent of side-chain hydroxylation of Ile- and Val-derived glucosinolates.

A single Danish plant (DK3a) among those analyzed was apparently deficient in accumulation of Ile-derived glucosinolates (**3**, **4**) and Phe/Tyr-derived **9**, but accumulated unusually high levels of Val-derived **2** and Phe-derived **7**, whereas a second plant (DK3b) from the same locality was also low in **9** but did contain Ile-derived glucosinolates (**3** and **4**) like other Danish plants (**Table 3**). As is often the case, substituted Trp-derived glucosinolates occurred mainly in roots. In roots of the commercially obtained accession the rare dimethoxy indole glucosinolate **14** was detected with certainty, whereas **14** was not detected in the remaining accessions. The incompletely identified iso**6** from roots of COM plants was not detected in roots or leaves from Danish accessions, but leaves from both MA1 and COM showed a minor peak with the correct retention time.

Relationship of Chromosome Numbers to Chemical Variation. To investigate whether the observed chemical variation was related to differences in ploidy level, we determined chromosome numbers in individual plants collected at the respective localities (**Figure 5**). However, the chemical variation within and between Danish and Massachusetts *C. pratensis* populations was not correlated with the ploidy or chromosome number as all of the individuals appeared to be $2n = 30$ (**Table 4**). Glucosinolate profiles of the sampled individuals confirmed that they were similar to the chemotypes previously collected at the same localities (**Table 4**). The aromatic **9** was not detectable in plant DK3c, suggesting that lack of this glucosinolate was common at locality DK3 and was not due to an alternative drying protocol for DK3a and DK3b.

DISCUSSION

We have discovered extensive variation in glucosinolate profiles among *C. pratensis* plants with chromosome number $2n = 30$ and identified a palette of 14 glucosinolates in the species, consisting of 5 structural groups apparently derived from the amino acids Val, Ile, dihomo-Ile, Phe or Tyr, and Trp (**Figure 2**).

The elucidation of the structure of the novel desulfoglucosinolate d**6** started with the observation of an apparent side-chain $-CH_2OH$ group, detected as a DEPT signal of a $-CH_2-$ group at 64.4 ppm, which is similar to chemical shifts of the thioglucose moiety $-CHOH-$ and $-CH_2OH$ signals (**Table 2**). The systematic assignment of NMR signals of d**6** started by identification of the two H atoms in position 1 by an HMBC interaction to the C=N carbon (which is regarded as position 0 according to the conventional numbering system of glucosinolates). From this position, assignment of the remaining signals was possible by COSY. The multiplicities and coupling patterns as revealed by COSY were compatible with the illustrated structure only (**Figure 1**; **Table 1**), including a single $-CH_3$ group and the branching $-CH_2-$ group already noted in the DEPT spectrum. From determination of elemental composition by high-resolution MS, the substituent on the branching $-CH_2-$ group was confirmed to be a hydroxy group. Assignment of carbon signals was straightforward by HSQC supported by DEPT and HMBC.

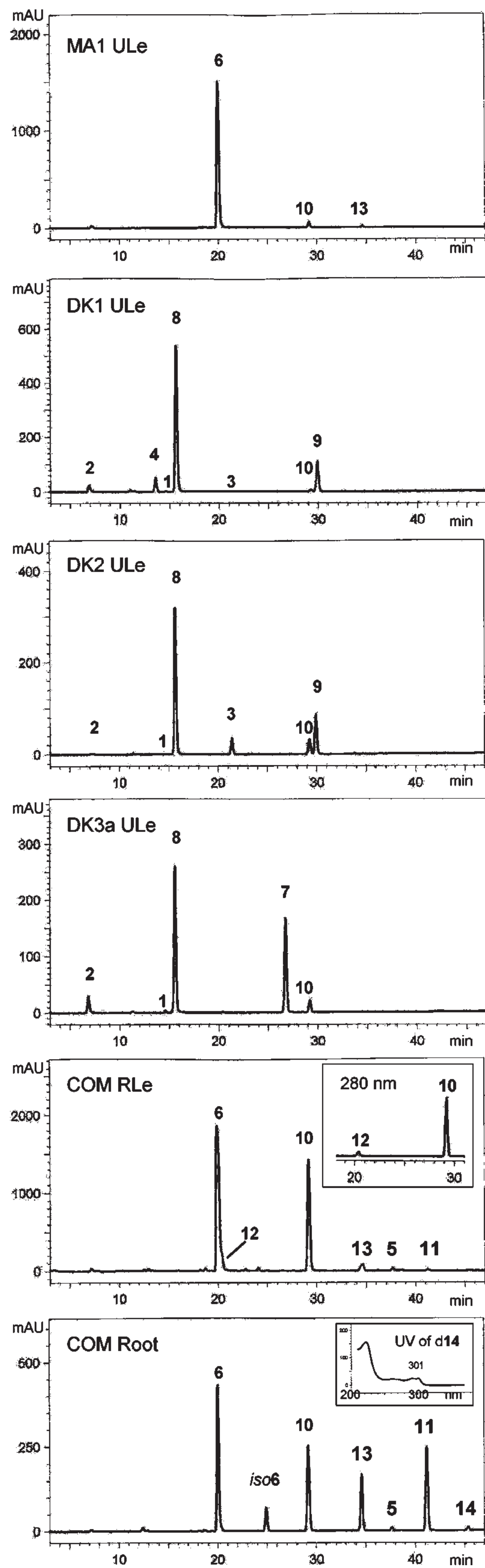

Figure 4. Demonstration of glucosinolate profile variation by HPLC-DAD analysis of desulfated derivatives from selected samples, as indicated by locality code and plant part designation (as in **Table 3**) in each panel. Desulfo glucosinolate peaks are labeled with the numbers of the corresponding intact glucosinolates for simplicity. UV detection was at 229 nm except for 280 nm in the box labeled so. The inserted UV spectrum of d**14** was measured after reinjection of the isolated, concentrated peak.

Table 3. Glucosinolate Profiles of Representative Individuals of *Cardamine pratensis* from Various Localities or of Commercial Origin (COM)[a]

origin:	MA1	DK1		DK2		DK3a		DK3b	COM		
plant part:	ULe	ULe	root	ULe	root	ULe	RLe	ULe	RLe	root	seed
glucosinolate											
1	nd	0.1	nd	0.1	0.3	0.4	1.3	0.4	nd	nd	nd
2	nd	**2.8**	**0.4**	0.2	0.1	**3.3**	**6.3**	0.1	nd	nd	nd
3	nd	0.1	nd	**3.3**	**1.5**	nd	nd	**10.7**	nd	nd	nd
4	nd	**5.5**	**0.4**	nd	nd	nd	nd	0.1	nd	nd	nd
5	0.3	nd	nd	tr	0.2	nd	nd	0.2	0.5	0.4	**4.7**
6	**82.8**	nd	nd	nd	nd	nd	nd	nd	**42.6**	**17.1**	**114.6**
iso**6**	tr	nd	nd	nd	nd	nd	nd	nd	0.1	**2.9**	0.2
7	nd	tr	nd	tr	nd	**18.0**	**35.3**	nd	nd	nd	nd
8	nd	**28.2**	**0.5**	**14.6**	**3.2**	**13.5**	**24.2**	**38.1**	nd	nd	nd
9	nd	**6.2**	**4.3**	**4.3**	**8.6**	nd	0.1	nd	nd	nd	nd
10	0.9	0.2	0.7	1.0	2.9	0.8	1.6	2.5	7.0	3.5	0.4
11	nd	nd	0.5	nd	tr	nd	nd	nd	tr	2.3	0.1
12	tr	nd	nd	nd	0.1	nd	0.1	tr	0.2	tr	nd
13	0.3	nd	0.2	nd	2.2	nd	nd	nd	tr	2.0	nd
14	nd	nd	nd	nd	nd	nd	nd	nd	nd	**0.1**	nd
not identified[b]	1.8	0.8	0.3	0.1	0.0	0.1	0.0	0.9	1.5	0.9	4.4
total	86.2	43.9	7.3	23.7	19.2	36.2	68.9	53.1	52.1	29.4	124.4

[a] Characteristic constituents are highlighted. Contents in μmol/g of dry wt. The profile of a single sample from each individual plant part is reported. ULe, upper leaf from flowering shoot; RLe, rosette leaf; nd, not detected; tr, traces. [b] Excluding iso**6**.

HMBC interactions between ^{13}C in the $-CH_2OH$ group and H-2, H-3, and H-4 but not H-1 and H-5 confirmed the position of the $-CH_2OH$ group already concluded from proton coupling patterns. The elucidated structure was also in accordance with the usual glucosinolate biosynthesis based on protein amino acids chain-elongated from the α-amino end (*1*); the structures of **6** and co-occurring **5** would be in accordance with this kind of biosynthesis from dihomo-Ile and represent an enzymatically simple modification of the biosynthesis of **3** and **4** from Ile as deduced for Danish plants. Hence, the side chains of **5** and **6** would be expected to retain the absolute configuration of Ile, but this expectation was not tested.

Lack of detection of lower homologues such as **3** and **4** in Massachusetts and commercial plants (**Table 3**) suggested that the glucosinolate biosynthetic enzymes in Massachusetts and commercial plants would not accept native Ile. On the other hand, detection of **5** in some Danish plants suggested that genes and enzymes with low activity for chain elongation of Ile were present in these plants. Considering the biochemical consequences of the variable chain elongation and hydroxylation in the species, it is interesting to speculate on the possibility of reaction between the isothiocyanate and alcohol functionalities in the corresponding isothiocyanates, but so far we have not investigated the hydrolysis products of **4** and **6**.

The detected low levels of the 1,4-dimethoxy indole glucosinolate **14** in roots from one accession was accompanied by high levels of both the corresponding 1-methoxy and 4-methoxy indole glucosinolates (**Figure 4**). This is in contrast to high levels of **14** and the *absence* of the 1-methoxy indole glucosinolate **11** in the previously characterized P-type of *Barbarea vulgaris* (*13*). The fact that **14** is rarely detectable at all in crucifers (*8*, *13*, *17*) is probably evidence of a general high specificity of the enzymes involved in the biosynthesis of **11** and **13**, and the trace of **14** observed in *C. pratensis* may simply be due to less than perfect specificity of a biosynthetic enzyme, rather than a new enzyme predicted in the P-type of *B. vulgaris* (*13*).

The only other recent report of glucosinolates in *C. pratensis* (*7*) lists a number of straight-chain and branched aliphatics with and

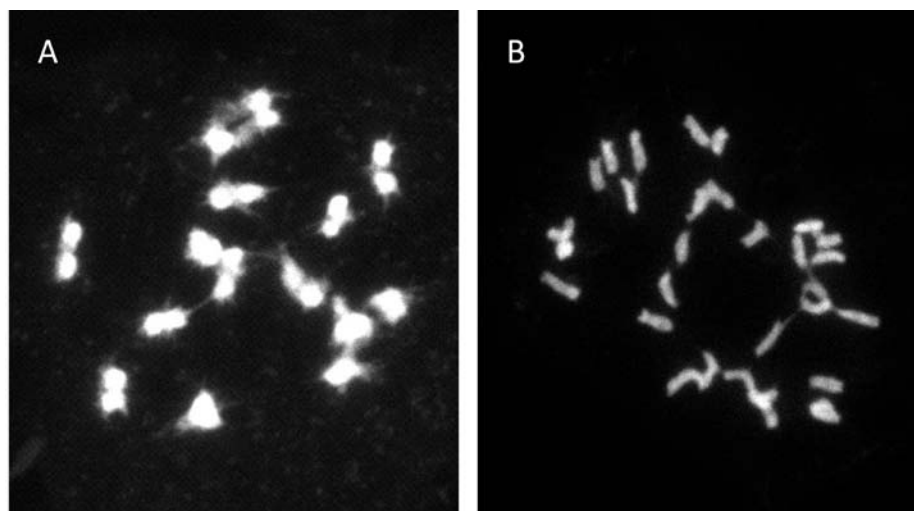

Figure 5. Stained chromosomes of *Cardamine pratensis*: (**A**) meiotic metaphase I pollen mother cell with 15 bivalents, 1 cm = 4.4 μm; (**B**) mitotic metaphase root-tip cell with 2*n* = 30 chromosomes, 1 cm = 7.7 μm.

Table 4. Chromosome Numbers in Pollen Mother Cells (*n*) and Root Tip Cells (2*n*) of *Cardamine pratensis* from the Four Investigated Localities and from Plants Grown from Commercial Seeds (COM), as well as Glucosinolate Profiles of Individual Plants Subjected to Chromosome Counts

	chromosome no.		dominating glucosinolates			
plant origin	*n*	2*n*	aliphatic (**1**−**6**)	Phe derived (**7**−**9**)	Trp derived (**10**−**14**)	plant part[a]
DK1	−[b]	c. 30	**4**, **2**	**8**	**10**,**13**	RLe[c]
DK2	15	30	**3**	**8**, **9**	**10**	RLe[c]
DK3(c)	15	30	**3**	**8**	**10**	RLe[c]
COM	−	30	**6**	nd[d]	**10**	RLe[c]
MA2[e]	15	30	**6**	nd	**10**	ULe[f]

[a] Plant part subjected to glucosinolate analysis. [b] Proper meiotic cells not obtained from this accession. [c] Rosette leaves from plants grown outdoors in pots at identical conditions from April to September. [d] nd, not detected. [e] Two individual plants tested with identical results. [f] Upper leaves of flowering plants grown at separate conditions.

without hydroxylation based on HPLC-MS analysis of seed extract. The structures identified by us in vegetative parts would in some cases have the same molecular weights, but did not in general have the same branching structures, and our results expand the reported range of structures from the species (*6*, *7*). We agree with the authors of the large-scale screening including *C. pratensis* seeds that confirmation of potentially uncertain peaks may require additional NMR experiments (*7*); the use of ^{1}H NMR for peak identification was possible with moderate analysis time from compounds isolated from 1 g of dry weight of pooled leaves even in the case of the relatively minor peaks d**1**−d**4** (**Figure 4**), and with rather excessive numbers of scans even trace peaks of known structures can be identified by ^{1}H NMR (*8*).

The discovery of a previously unknown structure also drew our attention to another potential limitation in HPLC-MS-based peak identification: our in-house routine peak annotation procedure initially assigned the peaks from **6** and iso**6** as potential 4-methylthiobut-3-enylglucosinolate, a known glucosinolate with the appropriate molecular mass and approximate retention time for the desulfated derivative (*22*). Accidental coelution might have led us to accept this identity, but the NMR investigation immediately excluded the presence of an unsaturated side chain. Even though the list of known glucosinolates appears to be relatively saturated (*6*−*11*), continued attention to the possibility of yet unknown structures and confirmation by NMR or comparable techniques is obviously still relevant. The apparent isomer of **6** in COM roots would seem to be an additional novel glucosinolate structure (*6*) and was present at levels that should allow identification by NMR (**Figure 3**), but this was unfortunately not accomplished because the harvested root sample was too small.

The identified glucosinolate structures are not in conflict with the majority of other investigations of the *Cardamine* genus. From seeds of *Cardamine flexuosa* With., Ala-derived **1**, Ile-derived **3**, and Phe-derived **7** and phenethylglucosinolate were reported (*2*), whereas seeds of *Cardamine hirsuta* L. were dominated by Met-derived but-3-enylglucosinolate and **7** and seeds of *Cardamine impatiens* L. and *Cardamine bulbifera* by but-3-enylglucosinolate (*2*, *7*). From rhizomes of *Cardamine diphylla* (Michx.) A. Woods, Val-derived **1**, Ile-derived **3**, homo-Ile-derived 2-methylbutylglucosinolate (dominating), dihomo-Ile-derived **5**, and Trp-derived **10**, **11**, and **13** were isolated and unequivocally identified (*23*), whereas foliage of the same species collected by us in Massachusetts was dominated by **1** according to HPLC-MS (unpublished results). An additional precursor amino acid for the genus, Leu, was deduced from the detection of high levels of 2-methylpropylglucosinolate in leaves of *Cardamine amara* (*24*). In contrast to the side chains devoid of hydroxy groups in the study above, an apparent polymorphism in hydroxylation of side-chain methyl groups of Val-, Ile-, and Leu-derived glucosinolates has been reported from *Cardamine cordifolia* A. Gray (*25*). Accumulation of **7** in plant DK3a low in *O*-methylated **9** (**Figure 4**) is evidence that *C. pratensis* may accumulate a Phe-derived glucosinolate like many other *Cardamine* species, so it may be that **8** and **9** in most Danish plants were also derived from Phe rather than Tyr. Apparently, there is substantial variation for chain elongation of Ile-derived glucosinolates and hydroxylation of Val-, Leu-, and Ile-derived glucosinolates in the genus *Cardamine*, as well as for glucosinolates derived from Met, Phe (and Tyr?), and Trp.

From the compilation above, there is also a tendency for two groups of species in the genus, with aliphatic glucosinolates derived from either Met or Val/Leu/Ile, although Met-derived glucosinolates in leaves were suggested to have been lost in a common ancestor of *Barbarea* and *Cardamine* (*24*). Glucosinolate side chains with no branching such as in *n*-butyl, *n*-pentyl, and *n*-hexyl glucosinolates have so far been reported from only one investigation of *Cardamine* (*7*) and may need confirmation, as identification was based on molecular mass and retention time only and as these natural products appear to be very rare in nature (*6*). Apparent Leu-derived glucosinolates such as 2-methylpropyl and 2-hydroxymethylpropyl glucosinolate (*7*) are not unprecedented in the genus (*24*, *25*), and some unidentified peaks in this investigation may be structurally related (**Figure 3**), but NMR would also in this case be helpful for distinguishing isomers.

C. pratensis is a species complex with a high degree of molecular, morphological, and cytological variation (*26*−*29*). The complex includes several widely distributed and sometimes intermixed cytotypes from diploids up to heptaploids. In Scandinavia, two ploidy levels are common: the tetraploid with $2n = 30$ and the octoploid with $2n = 56$ (56−64) (*29*). Molecular investigations of closely related species of *Cardamine* including the various cytotypes of *C. pratensis* were not able to recognize these *C. pratensis* cytotypes as separate taxa (*26*) and concluded that these taxa made up a continuum of intermixed forms.

The molecular phylogenetic investigations on the genus do not include material of *C. pratensis* from North America (*26*−*28*). The shared opinion among *Cardamine* scientists (*26*, *28*) is that it is a fairly young genus having its center of diversity in Eurasia. After the Ice Ages, it went through series of polyploidization and speciation events and was dispersed, probably by animals, to North America. From demonstration of a complex phylogeny of diploid taxa from the *C. pratensis* species complex, it was suggested that the phylogeny of polyploid taxa could be even more complex (*27*). Our finding of identical chromosome numbers of *C. pratensis* plants with contrasting glucosinolate profiles would seem to support this suggestion and provides an example of a variable character of biological significance (*3*) in the species complex. As the Danish and commercially obtained plants (of uncertain geographical origin but with the same glucosinolate profile as detected in Massachusetts plants) were cocultivated for an extended period before sampling, it is likely that variation in glucosinolate profiles had a genetic basis. The degree of sampling achieved here is obviously totally insufficient for characterization of populations on two entire continents. If the distinct differences between the low number of southern Scandinavian and eastern North American accessions investigated here are representative, glucosinolates may have been subject to evolution after dispersal to North America (*24*, *26*, *28*), but if the investigated commercial seeds were indeed of German origin, the picture may be more complex. However, our results suggest the relevance of including Scandinavian and North American material in future phylogenetic investigations.

The considerable levels of the novel glucosinolate **6** in the species should also facilitate general investigations of its hydrolysis products and of structure-dependent bioactivity of glucosinolates. Furthermore, the considerable biological variation in *C. pratensis* plants with identical chromosome number may allow genetic and ecological investigations of branched-chain and aromatic glucosinolate activities. This structurally unique variation complements the natural variation currently investigated in other species (*13*−*16*). The presence of high levels of either the aromatic **8** or aliphatic **6** is also an interesting complement of the variation of **8** obtained by genetic modification (*3*) and resembles the loss of **8** and gain of branched-chain aliphatics (**3**, **5**) deduced in the evolution of some plant species related to *Sinapis alba* (*17*), whereas the tendency for either Met- or branched-chain amino acid-derived glucosinolates in *Cardamine* resembles a polymorphism linked to differential insect susceptibility in *Boechera stricta* (*15*). *C. pratensis* is reported as a host plant for butterflies with specialized feeding habits, with *Anthocharis cardamines* consuming mainly floral parts (see, e.g., ref *29*) and *Pieris brassicae* consuming foliage in early stages and floral parts in later stages (*30*). In separate publications, we report our investigations of plant-part preferences and plant chemistry in these plant−herbivore interactions (unpublished results) and of the involvement of *O*-demethylase activity in insect metabolism of methoxyaryl glucosinolates (*20*) such as **9** (unpublished results).

ACKNOWLEDGMENT

We thank Birgitte B. Rasmussen for skillful glucosinolate isolation and analysis, Ib Linde-Laursen and Karen Rysbjerg Munk for their skillful assistance in chromosome analysis, the staff at Chiltern Seeds and Jelitto Perennial Seeds for kindly providing information on seed origin, and the editor and three anonymous reviewers for helpful suggestions to an earlier version of the text.

LITERATURE CITED

(1) Halkier, B. A.; Gershenzon, J. Biology and biochemistry of glucosinolates. *Annu. Rev. Plant Biol.* **2006**, *57*, 303–333.
(2) Daxenbichler, M. E.; Spencer, G. F.; Carlson, D. G.; Rose, G. B.; Brinker, A. M.; Powell, R. G. Glucosinolate composition of 297 species of wild plants. *Phytochemistry* **1991**, *30*, 2623–2638.
(3) Brader, G.; Mikkelsen, M. D.; Halkier, B. A.; Palva, E. T. Altering glucosinolate profiles modulates disease resistance in plants. *Plant J.* **2006**, *46*, 758–767.
(4) Mumm, R.; Burow, M.; Bukovinszkine'Kiss, G.; Kazantzidou, E.; Wittstock, U.; Dicke, M.; Gershenzon, J. Formation of simple nitriles upon glucosinolate hydrolysis affects direct and indirect defense against the specialist herbivore, *Pieris rapae*. *J. Chem. Ecol.* **2008**, *34*, 1311–1321.
(5) Clay, N. K.; Adio, A. M.; Denoux, C.; Jander, G.; Ausubel, F. M. Glucosinolate metabolites required for an *Arabidopsis* innate immune response. *Science* **2009**, *323*, 95–101.
(6) Fahey, J. W.; Zalcmann, A. T.; Talalay, P. The chemical diversity and distribution of glucosinolates and isothiocyanates among plants. *Phytochemistry* **2001**, *56*, 5–51.
(7) Bennett, R.; Mellon, F. A.; Kroon, P. A. 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.* **2004**, *52*, 428–438.
(8) Agerbirk, N.; Petersen, B. L.; Olsen, C. E.; Halkier, B. A.; Nielsen, J. K. 1,4-Dimethoxyglucobrassicin in *Barbarea* and 4-hydroxyglucobrassicin in *Arabidopsis* and *Brassica*. *J. Agric. Food Chem.* **2001**, *49*, 1502–1507.
(9) Frechard, A.; Fabre, N.; Pean, C.; Montaut, S.; Fauvel, M.-T.; Rollin, P.; Fourasté, I. Novel indole-type glucosinolates from woad (*Isatis tinctoria* L.). *Tetrahedron Lett.* **2001**, *42*, 9015–9017.
(10) Bennett, R. N.; Mellon, F. A.; Botting, N. P.; Eagels, J.; Rosa, E. A. S.; Williamson, G. Identification of the major glucosinolate (4-mercaptobutyl glucosinolate) in leaves of *Eruca sativa* L. (salad rocket). *Phytochemistry* **2002**, *61*, 25–30.
(11) Reichelt, M.; Brown, P. D.; Schneider, B.; Oldham, N. J.; Stauber, E.; Tokuhisa, J.; Kliebenstein, D. J.; Mitchell-Olds, T.; Gershenzon, J. Benzoic acid glucosinolate esters and other glucosinolates from *Arabidopsis thaliana*. *Phytochemistry* **2002**, *59*, 663–671.
(12) Bidart-Bouzat, M. G.; Kliebenstein, D. J. Differential levels of insect herbivory in the field associated with genotypic variation in glucosinolates in *Arabidopsis thaliana*. *J. Chem. Ecol.* **2008**, *34*, 1026–1037.

(13) Agerbirk, N.; Ørgaard, M.; Nielsen, J. K. Glucosinolates, flea beetle resistance, and leaf pubescence as taxonomic markers in the genus *Barbarea* (Brassicaceae). *Phytochemistry* **2003**, *63*, 69–80.

(14) van Leur, H.; Vet, L. E. M.; van der Putten, W. H.; van Dam, N. M. *Barbarea vulgaris* glucosinolate phenotypes differentially affect performance and preference of two different species of lepidopteran herbivores. *J. Chem. Ecol.* **2008**, *34*, 121–131.

(15) Schranz, M. E.; Manzaneda, A. J.; Windsor, A. J.; Clauss, M. J.; Mitchell-Olds, T. Ecological genomics of *Boechera stricta*: identification of a QTL controlling the allocation of methionine- vs branched-chain amino acid-derived glucosinolates and levels of insect herbivory. *Heredity* **2009**, *102*, 465–474.

(16) Gols, R.; Wagenaar, R.; Bukovinszky, T.; van Dam, N. M.; Dicke, M.; Bullock, J. M.; Harvey, J. A. Genetic variation in defense chemistry in wild cabbages affect herbivores and their endoparasitoids. *Ecology* **2008**, *89*, 1616–1626.

(17) Agerbirk, N.; Warwick, S.; Hansen, P. R.; Olsen, C. E. *Sinapis* phylogeny and evolution of glucosinolates and specific nitrile degrading enzymes. *Phytochemistry* **2008**, *69*, 2937–2949.

(18) Keeler, M. S.; Chew, F. S. Escaping an evolutionary trap: preference and performance of a native insect on an exotic invasive host. *Oecologia* **2008**, *156*, 559–568.

(19) Wathelet, J.-P.; Iori, R.; Leoni, O.; Rollin, P.; Quinsac, A.; Palmieri, S. Guidelines for glucosinolate analysis in green tissues used for biofumigation. *Agroindustria* **2004**, *3*, 257–266.

(20) Agerbirk, N.; Olsen, C. E.; Poulsen, E.; Jacobsen, N.; Hansen, P. R. Complex metabolism of aromatic glucosinolates in *Pieris rapae* caterpillars involving nitrile formation, hydroxylation, demethylation, sulfation, and host plant dependent carboxylic acid formation. *Insect Biochem. Mol. Biol.* **2010**, *40*, 126–137.

(21) Ørgaard, M.; Jacobsen, N.; Heslop-Harrison, J. S. The hybrid origin of two cultivars of *Crocus* (Iridaceae) analyzed by molecular cytogenetics including genomic Southern and *in situ* hybridization. *Ann. Bot.* **1995**, *76*, 253–262.

(22) Barillari, J; Cervellati, R.; Paolini, M.; Tatibouët, A.; Rollin, P.; Iori, R. Isolation of 4-methylthio-3-butenyl glucosinolate from *Raphanus sativus* sprouts (kaiware daikon) and its redox properties. *J. Agric. Food Chem.* **2005**, *53*, 9890–9896.

(23) Montaut, S.; Bleeker, R. S.; Jacques, C. Phytochemical constituents of *Cardamine diphylla*. *Can. J. Chem.* **2010**, *88*, 50–55.

(24) Windsor, A. J.; Reichelt, M.; Figuth, A.; Svatoš, A.; Kroymann, J.; Kliebenstein, D. J.; Gershenzon, J.; Mitchell-Olds, T. Geographic and evolutionary diversification of glucosinolates among near relatives of *Arabidopsis thaliana* (Brassicaceae). *Phytochemistry* **2005**, *66*, 1321–1333.

(25) Rodman, J. E.; Chew, F. S. Phytochemical correlates of herbivory in a community of native and naturalized Cruciferae. *Biochem. Syst. Ecol.* **1980**, *8*, 43–50.

(26) Franzke, A.; Hurka, H. Molecular systematics and biogeography of the *Cardamine pratensis* complex (Brassicaceae). *Plant Syst. Evol.* **2000**, *224*, 213–234.

(27) Marhold, K.; Lihova, J.; Perny, M.; Bleeker, W. Comparative ITS and AFLP analysis of diploid *Cardamine* (Brassicaceae) taxa from closely related polyploid complexes. *Ann. Bot.* **2004**, *93*, 507–520.

(28) Carlsen, T.; Bleeker, W.; Hurka, H.; Elven, R.; Brochmann, C. Biogeography and phylogeny of *Cardamine* (Brassicaceae). *Ann. Mo. Bot. Gard.* **2009**, *96*, 215–236.

(29) Arvanitis, L.; Wiklund, C.; Ehrlén, J. Plant ploidy level influences selection by butterfly seed predators. *Oikos* **2008**, *117*, 1020–1025.

(30) Smallegange, R. C.; van Loon, J. J. A.; Blatt, S. E.; Harvey, J. A.; Agerbirk, N.; Dicke, M. Flower vs. leaf feeding by *Pieris brassicae*: glucosinolate-rich flower tissues are preferred and sustain higher growth rate. *J. Chem. Ecol.* **2007**, *33*, 1831–1844.

Received for review December 9, 2009. Revised manuscript received March 11, 2010. Accepted March 11, 2010. This work was supported financially by the Torben og Alice Frimodts Fond.