Available online at www.sciencedirect.com
SCIENCE DIRECT®
Phytochemistry 66 (2005) 1321–1333
PHYTOCHEMISTRY
www.elsevier.com/locate/phytochem

Geographic and evolutionary diversification of glucosinolates among near relatives of *Arabidopsis thaliana* (Brassicaceae)

Aaron J. Windsor, Michael Reichelt, Antje Figuth, Aleš Svatoš, Juergen Kroymann, Daniel J. Kliebenstein [1], Jonathan Gershenzon, Thomas Mitchell-Olds *

Max-Planck-Institute for Chemical Ecology, Genetics and Evolution, Hans-Knoell-Strasse 8, D-07745 Jena, Germany

Received 13 December 2004; received in revised form 21 January 2005
Available online 23 May 2005

Abstract

Glucosinolates are biologically active secondary metabolites that display both intra- and interspecific variation in the order Brassicales. Glucosinolate profiles have not been interpreted within a phylogenic framework and little is known regarding the processes that influence the evolution of glucosinolate diversity at a macroevolutionary scale. We have analyzed leaf glucosinolate profiles from members of the Brassicaceae that have diverged from *Arabidopsis thaliana* within the last 15 million years and interpreted our findings relative to the phylogeny of this group. We identified several interspecific polymorphisms in glucosinolate composition. A majority of these polymorphisms are lineage-specific secondary losses of glucosinolate characters, but a gain-of-character polymorphism was also detected. The genetic basis of most observed polymorphisms appears to be regulatory. In the case of *A. lyrata*, geographic distribution is also shown to contribute to glucosinolate metabolic diversity. Further, we observed evidence of gene-flow between sympatric species, parallel evolution, and the existence of genetic constraints on the evolution of glucosinolates within the Brassicaceae.

Keywords: Brassicaceae; *Arabidopsis lyrata*; Glucosinolate; Secondary metabolism; Genetic constraint; Evolution

1. Introduction

1.1. Phylogenetics and glucosinolate evolution

Glucosinolates, or mustard oil glycosides, are biologically active secondary metabolites found in the Brassicaceae and related families (Kjaer and Schuster, 1972; Kjaer, 1976; Doughty et al., 1991; Rodman et al., 1996, 1998; Raybould and Moyes, 2001; Fahey et al., 2001; Tokuhisa et al., 2004). These compounds are genetically variable within plant species, and strongly influence the feeding choices of insect herbivores (Lambrix et al., 2001; Kroymann et al., 2003). In *Arabidopsis thaliana*, genetic polymorphisms and loss-of-function mutations have allowed identification of genes encoding glucosinolate biosynthetic enzymes (Kliebenstein et al., 2001a,b; Hemm et al., 2003; Reintanz et al., 2001; Kroymann et al., 2001), and these results have been extended to *Brassica* relatives of *A. thaliana* (Gao et al., 2004; Li and Quiros, 2003; Quiros et al., 2001).

On a macroevolutionary time-scale, comparative phylogenetic analyses can elucidate biological patterns and processes influencing the evolution of secondary metabolism over millions of years. Although several compendia have cataloged glucosinolate profiles of tens or hundreds of species (Fahey et al., 2001; Daxenbichler et al., 1991), previous studies have not included a

* Corresponding author. Tel.: +49 3641 571401; fax: +49 3641 571402.
E-mail address: tmo@ice.mpg.de (T. Mitchell-Olds).
[1] Current address: Department of Vegetable Crops, University of California, Davis, CA 95616, USA.

0031-9422/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.phytochem.2005.04.016

detailed phylogenetic framework to interpret similarities and differences in glucosinolate patterns. Here we examine evolutionary changes in glucosinolate metabolism among close relatives of *A. thaliana*, which have diverged within the last 15 million years (Koch et al., 2001), to address evolutionary hypotheses and processes among species. In the following study, we have generated glucosinolate profiles for relatives of *A. thaliana* and used these data in a comparative framework to ask the following questions: (1) Are glucosinolate polymorphisms within *A. thaliana* also found in related species? (2) Do differences in glucosinolate profiles between species and within species involve changes to the structural portion of enzyme-encoding loci, or regulatory functions which impact gene expression? (3) Is there evidence for genetic constraints which limit the potential for future adaptation of glucosinolate profiles in entire lineages?

1.2. Biology of glucosinolates

Glucosinolates are synthesized from a limited repertoire of amino acids (Fig. 1) with a core structure consisting of an (*Z*)-*N*-hydroximinosulfate ester, a β-thio-linked glucose moiety, and a variable amino acid derived R-group (Fig. 1). In the chemical defense of plants against insect herbivores, glucosinolates function with a specific class of β-thioglucosidases, the myrosinases. Glucosinolates and myrosinase coexist temporally in intact plant tissues, but are separated spatially (Eriksson et al., 2001; Husebye et al., 2002; Geshi et al., 1998; Thangstad et al., 2004). These two elements come into contact when tissue damage occurs, leading to cleavage of the glucosinolate thioglucoside linkages, thereby liberating D-glucose and a thiohydroximate-*O*-sulfate. Thiohydroximate-*O*-sulfates are labile and undergo spontaneous rearrangements to form a variety of toxic products including nitriles, isothiocyanates, and thiocyanates. It is the quantity and spectrum (i.e. quality) of the end-products of the glucosinolate–myrosinase interaction that confer the biological activity associated with glucosinolates; the end-products defend against generalist insect herbivores, or can act as ovipositional cues and feeding stimulants for specialist insect herbivores (Raybould and Moyes, 2001; Pivnick et al., 1994; Marazzi and Staedler, 2004). Several intraspecific studies have documented genetic variation for glucosinolate type and concentration (Kjaer, 1976; Louda and Rodman, 1983; Kliebenstein et al., 2001a; Fahey et al., 2001; Kroymann et al., 2003; Castro et al., 2004; Font et al., 2004; Charron and Sams, 2004) and demonstrated their functional significance in resistance to generalist insect herbivores (Raybould and Moyes, 2001; Kroymann et al., 2003; Kroymann and Mitchell-Olds, 2004). These results, in plant species separated by tens of millions of years of independent evolution (Koch et al., 2001), underscore the ecological importance of these secondary metabolites.

While structural variation among glucosinolates is dependent on the precursor amino acids, modifications of the R-groups also contribute enormous diversity to this class of compounds. Carbon chain-elongation of methionine (Fig. 2; Kroymann et al., 2001; Field et al., 2004; Textor et al., 2004) prior to the core pathway of glucosinolate biosynthesis is a common and well characterized example of side chain elaboration. After core biosynthesis, chemical modification of R-groups can occur through oxidation (Fig. 2), hydroxylation, and

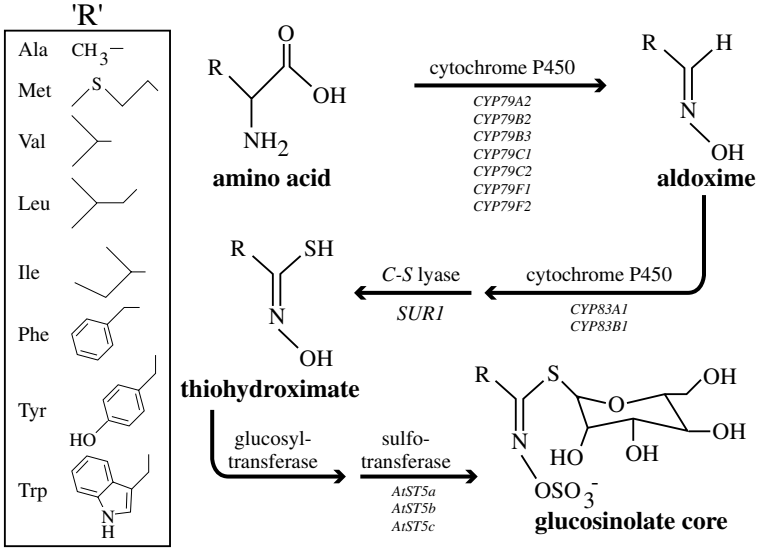

Fig. 1. The core pathway for biosynthesis of glucosinolates from amino acids. R-groups for each of the eight precursor amino acids are indicated. Loci from *A. thaliana* which encode the first three steps are indicated. Recently, the glucosinolate *C-S* lyase has been identified as the protein product of the *SUPERROOT1* (*SUR1*) locus in *A. thaliana* (Mikkelsen et al., 2004).

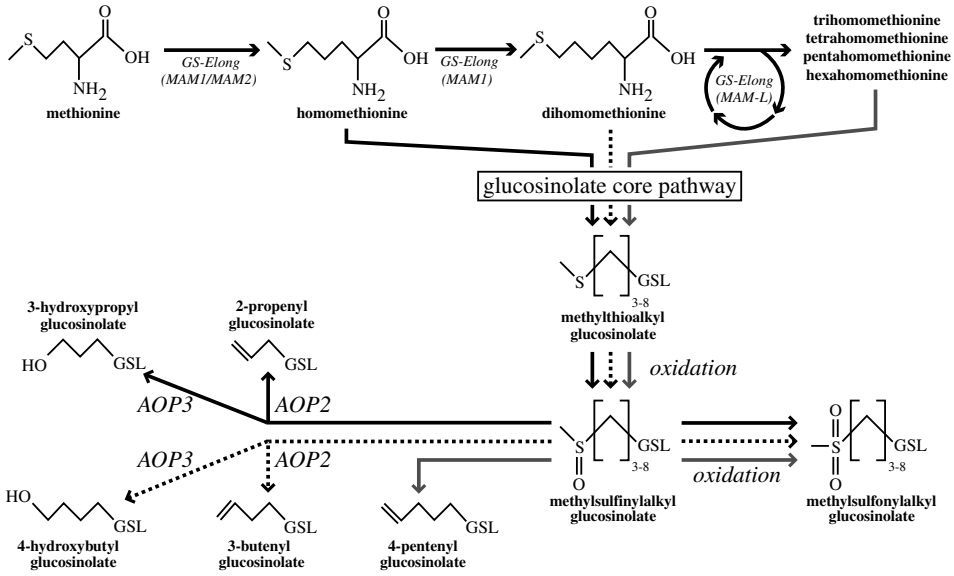

Fig. 2. Examples of side chain modification in methionine-derived glucosinolates. Where applicable, *A. thaliana* loci encoding specific enzymatic functions are indicated. 'GSL' indicates the glucosinolate core structure as presented in Fig. 1. Prior to core glucosinolate formation, methionine undergoes a series of elongation cycles yielding derivatives with three (homomethionine) to eight (hexahomomethionine) methylene groups in the side chain. The predominant side chain length is determined by the configuration of the *GS-Elong* complex locus. Glucosinolate assembly (Fig. 1) produces methylthioalkyl glucosinolates which, in turn, may be oxidized to methylsulfinylalkyl glucosinolates. Further oxidation produces methylsulfonylalkyl glucosinolates, but this chemical species has not been observed in *A. thaliana*. The methylsulfinyl moiety of methylsulfinylalkyl glucosinolates is cleaved by the products of *AOP2* or *AOP3* (Kliebenstein et al., 2001b), components of the *ALK-OHP* complex locus, to yield hydroxylalkyl or alkenyl glucosinolates, respectively. In *A. thaliana*, the reactions catalyzed by AOP2 and AOP3 appear to be restricted to 3-MSOP and 4-MSOB glucosinolates, however, 4-pentenyl glucosinolate has been observed in this species (Kliebenstein et al., 2001a).

esterification (Graser et al., 2001; reviewed in Tokuhisa et al., 2004). To date, greater than 120 glucosinolates have been identified within the order Brassicales (Fahey et al., 2001), with *A. thaliana* alone possessing more than 30 glucosinolate variants (Kliebenstein et al., 2001a; Reichelt et al., 2002; Brown et al., 2003).

Gene duplication is an important source of glucosinolate metabolic diversity in *A. thaliana*. The first reaction in the synthesis of glucosinolates, the conversion of a precursor amino acid to the aldoxime (Fig. 1), is catalyzed by the *CYP79*-family of cytochrome P450s. Seven expressed *CYP79* paralogs are present in the genome of *A. thaliana* ecotype Columbia (Mikkelsen et al., 2002). Individual CYP79s display distinct substrate specificities (Wittstock and Halkier, 2000; Hull et al., 2000; Mikkelsen et al., 2000; Chen et al., 2003; Reintanz et al., 2001) and differential expression in response to developmental cues and defense signaling pathways (Chen et al., 2003; Mikkelsen et al., 2003). Similarly, the *CYP83*-family P450s, which catalyze the second step in glucosinolate core biosynthesis (Fig. 1), are divergent with respect to substrate specificities and catalytic efficiencies (Bak and Feyereisen, 2001; Naur et al., 2003).

The allelic state of the *GS-Elong* locus determines the distribution of carbon-chain lengths observed among methionine-derived glucosinolates (Met-glucosinolates, Fig. 2; Kroymann et al., 2001; Field et al., 2004). In *A. thaliana*, the locus is complex, containing two or three directly repeated paralogous sequences (*MAM1*, *MAM2*, and *MAM-L*) which encode methylthioalkylmalate synthases (Kroymann et al., 2003). Individuals in whom the predominant Met-glucosinolates have side chains composed of three methylene groups (3C; one round of chain-elongation, Fig. 2) have a functional *MAM2* component, but probably lack MAM1 activity. Individuals with primarily four methylene groups in their glucosinolate side chains (4C; two rounds of chain-elongation, Fig. 2) express *MAM1*, which is epistatic to *MAM2* (Kroymann et al., 2001, 2003). MAM-L activity results in low levels of longer chain glucosinolates in both 3C and 4C individuals (Field et al., 2004). *GS-Elong* is highly polymorphic within *A. thaliana*, where insertion/deletion events and gene conversion contributing substantially to the evolution of the locus and, hence, the diversity of glucosinolate metabolism (Kroymann et al., 2003).

ALK-OHP is another complex locus that affects glucosinolate metabolic diversity by encoding the activities necessary for the conversion of methylsulfinylalkyl glucosinolates to either hydroxyalkyl or alkenyl glucosinolates (Fig. 2; Kliebenstein et al., 2001b; Li and Quiros, 2003). In *A. thaliana*, the locus is comprised of four tandemly duplicated sequences encoding 2-oxoglutarate-dependent dioxygenases (*AOP1*, *AOP-Ψ*, *AOP2*, and *AOP3*). *AOP1* is expressed ubiquitously among *A. thaliana* accessions; AOP2 and AOP3 activities are restricted to accessions producing alkenyl or hydroxyalkyl glucosinolates, respectively. The various

ALK-OHP alleles have not been studied as extensively as those of *GS-Elong*. However, the structural similarities between the two loci suggest similar potential to influence the evolution of glucosinolate metabolic diversity.

2. Results and discussion

2.1. Glucosinolate variation: The core pathway

Presence or absence is the most fundamental type of phenotypic polymorphism for any metabolic character. Within the phylogeny presented in Fig. 3(a), every clade produces glucosinolates of some form or another in the leaves (Tables 1 and 2), except the clade formed by *Arabis pendula*, *Camelina microcarpa*, and *Neslia paniculata* (the 'Neslia group', Fig. 3(a)) which produces no glucosinolates in leaves. The Neslia group is nested within glucosinolate producing clades (Table 1 and Fig. 3(a)) implying that the absence of glucosinolates represents a single secondary loss of the character in this lineage. It may be that the observed change results from an alteration in a major regulatory element or blockage in the core biosynthetic pathway after the reaction catalyzed by the *CYP83*-family of enzymes (Fig. 1), as all glucosinolates, and not a subset of structures, are missing. Further, the mutation is likely to be regulatory in nature, and not structural, as methylsulfinylalkyl glucosinolates have been identified in the seeds of *Neslia paniculata* (Kjaer and Schuster, 1972; Daxenbichler et al., 1991) and *Camelina microcarpa* (Daxenbichler et al., 1991), demonstrating that the core biosynthetic pathway is intact in these species.

2.1.1. Variation in Met-glucosinolates

Met-glucosinolates are the most common and abundant glucosinolates in the leaves of *A. thaliana* (Kliebenstein et al., 2001a; Brown et al., 2003) and they are present in the leaves of the majority of taxa presented in Fig. 3(a). Met-glucosinolates are, however, conspicuously absent from *Barbarea intermedia*, *B. verna*, and *Cardamine amara* (Table 1). These species form a clade, the 'Cardamine alliance' (Mitchell-Olds et al., 2004), together with the Met-glucosinolate containing species (Table 1) *Rorippa islandica* (Fig. 3(a)). Met-glucosinolate loss in *Barbarea* and *Cardamine* appears to be another example of secondary loss, as the clade is nested within lineages that produce Met-glucosinolates (Table 1 and Fig. 3(a)). Two equally parsimonious explanations exist for the pattern of Met-glucosinolate presence/absence observed in the Cardamine alliance. First, the capacity to synthesize Met-glucosinolates may have been lost a single time in the lineage leading to the most recent common ancestor (MRCA) of *Barbarea*, *Rorippa*, and *Cardamine* (Fig. 3(a)) and with *Rorippa* later reverting to production of Met-glucosinolates in leaves after divergence from the MRCA with *Barbarea*. Alternatively, the capacity to synthesize Met-glucosinolates may have been lost twice; once after the divergence of *Cardamine* from the MRCA with *Barbarea* and *Rorippa* and a second time after the divergence of *Barbarea* from the MRCA shared with *Rorippa*. We are inclined towards the first hypothesis – while *Rorippa islandica* produces a Met-glucosinolate, the observed profile is drastically different from that observed in other species in terms of both quantity and quality (Table 1). This is suggestive of an incomplete reversion to a full, Met-glucosinolate phenotype.

The members of the Cardamine alliance produce other, non-Met-derived glucosinolates (Table 2), raising the possibility that the clade is deficient in a Met-specific element of the biosynthetic pathway. This argues that, as observed in *A. thaliana*, components of the core glucosinolate biosynthetic pathway are partly modular, i.e. components are non-redundant and catalyze equivalent reactions while displaying varying substrate specificities. In *A. thaliana*, the conversion of Met-derivative amino acids to aldoximes (Fig. 1) is catalyzed by CYP79F1 and CYP79F2 (Reintanz et al., 2001; Hansen et al., 2001; Chen et al., 2003) with the resultant Met-aldoximes being metabolized most efficiently by the product of *CYP83A1* (Fig. 1; Naur et al., 2003). Even the final reaction in glucosinolate biosynthesis, the conversion of *desulfo*-glucosinolates to the core glucosinolate structure is catalyzed by a family of sulfotransferases that display some level of substrate specificity (Fig. 1; Piotrowski et al., 2004). Thus, lesions involving *CYP79Fx*, *CYP83A1*, or *AtST5x* homologs could be expected to cause the Met-glucosinolate phenotypes observed within the Cardamine alliance.

2.1.2. Variation in Met-glucosinolates: Chain elongation

Within *A. thaliana*, accessions can be divided into two phenotypic classes with regard to carbon-chain elongation: 3C and 4C (Kliebenstein et al., 2001a; Kroymann et al., 2003). While the major Met-glucosinolate component(s) of each phenotypic class have undergone one or two rounds of chain elongation, respectively, this species also produces longer chain length Met-glucosinolates (Kliebenstein et al., 2001a; Kroymann et al., 2001). The other members of the genus *Arabidopsis* exhibit similar, but distinct trends (Table 1). For example, all *A. lyrata* accessions assayed in this experiment produce mainly 3C Met-glucosinolates and lower levels of longer chain glucosinolates. *A. halleri* and *A. arenosa*, however, produce primarily long-chain glucosinolates (5C and greater) with extremely low or undetectable levels of 3C and 4C glucosinolates, possibly suggesting improved efficiency of the MAM-L-like activity in these species.

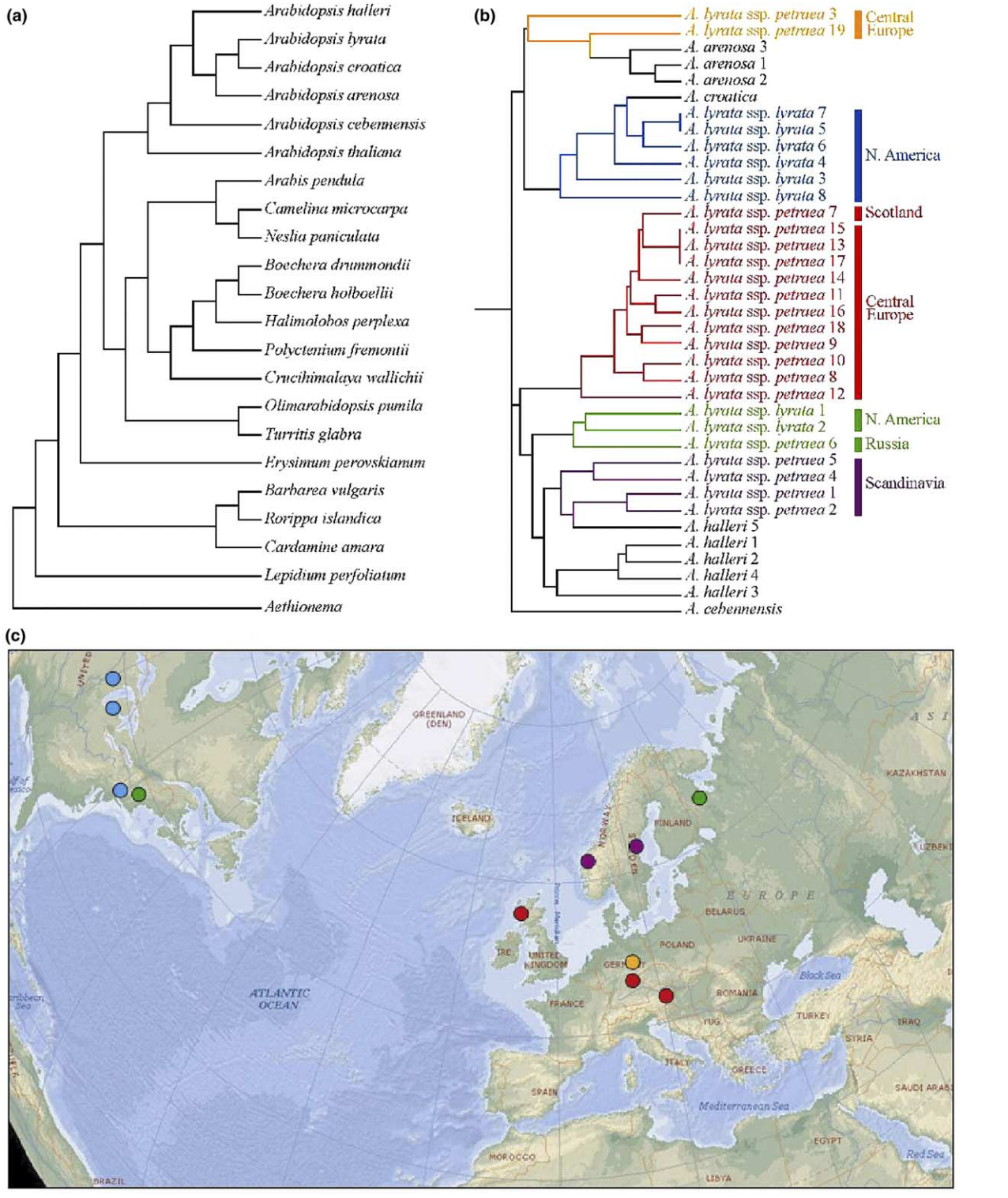

Fig. 3. Geographic and evolutionary diversification of glucosinolate metabolism in the Brassicaceae. (a) Phylogeny of the genus *Arabidopsis* and near relatives. The cladogram is based on Koch et al. (2001) and Al-Shehbaz et al. (2002), and supplemented with additional, unpublished data. With the exceptions of *Arabidopsis* and *Boechera*, a single exemplar species is provided for each genus. (b) Cluster analysis of glucosinolate metabolism for the genus *Arabidopsis*. For *A. lyrata*, the geographic distribution of contemporary phenotypes is indicated and color-coded. (c) Locations of Pleistocene refugia as suggested by the cluster analysis depicted in panel 'B'; color-coding is consistent with panel 'B'.

Outside of the genus *Arabidopsis*, the phenotypic patterns of Met-elongation are preserved with the exception of *Erysimum*. No long chain Met-glucosinolates are detected in this genus (Table 1) and the phylogenetic position of *Erysimum* (Fig. 3(a)) is suggestive of a single, secondary loss of MAM-L activity-like.

Taken together, the Met chain elongation data presented here indicate that MAM1, MAM2, and MAM-L activities are not restricted to *A. thaliana* and that they are separable and distinct in other species and genera. Further, MAM-L activity is generally observed in conjunction with one of the other two Met-elongation activities. Thus, the complex nature of *GS-Elong* is likely to be a feature found throughout the Brassicaceae.

Table 1
Distribution of methionine-derived glucosinolates in the genus *Arabidopsis* and other Brassicaceae

Species	AA[*]	*GS-Elong*[a]	3OHP	3MSOP	3MSOOP	3MTP	Allyl	4MSOB	4MSOOB	3-But
Arabidopsis thaliana[b]	Met	3C/4C	+	+	–	+	+	+	–	+
Arabidopsis lyrata ssp. *lyrata* 1	Met	3C	–	2	–	3	–	–	–	–
Arabidopsis lyrata ssp. *lyrata* 2	Met	3C	–	3	–	–	–	–	–	–
Arabidopsis lyrata ssp. *lyrata* 3	Met	3C	–	–	–	–	3	–	–	–
Arabidopsis lyrata ssp. *lyrata* 4	Met	3C	–	1	–	–	4	–	–	–
Arabidopsis lyrata ssp. *lyrata* 5	Met	3C	–	–	–	–	4	–	–	–
Arabidopsis lyrata ssp. *lyrata* 6	Met	3C	–	–	–	–	4	–	–	–
Arabidopsis lyrata ssp. *lyrata* 7	Met	3C	–	–	–	–	4	–	–	–
Arabidopsis lyrata ssp. *lyrata* 8	Met	3C	1	2	–	1	4	–	–	–
Arabidopsis lyrata ssp. *petraea* 1	Met	3C	–	–	2	–	–	–	1	–
Arabidopsis lyrata ssp. *petraea* 2	Met	3C	–	–	1	–	–	–	–	–
Arabidopsis lyrata ssp. *petraea* 3	Met	3C	–	–	4	–	–	–	–	–
Arabidopsis lyrata ssp. *petraea* 4	Met	3C	–	2	–	1	–	–	–	–
Arabidopsis lyrata ssp. *petraea* 5	Met	3C	3	1	–	1	–	–	–	–
Arabidopsis lyrata ssp. *petraea* 6	Met	3C	1	2	–	1	1	–	–	–
Arabidopsis lyrata ssp. *petraea* 7	Met	3C	4	–	–	–	–	–	–	–
Arabidopsis lyrata ssp. *petraea* 8	Met	3C	3	–	–	–	–	–	–	–
Arabidopsis lyrata ssp. *petraea* 9	Met	3C	3	–	–	–	–	–	–	–
Arabidopsis lyrata ssp. *petraea* 10	Met	3C	4	–	–	–	–	–	–	–
Arabidopsis lyrata ssp. *petraea* 11	Met	3C	4	–	–	–	–	–	–	–
Arabidopsis lyrata ssp. *petraea* 12	Met	3C	4	1	–	1	1	–	–	–
Arabidopsis lyrata ssp. *petraea* 13	Met	3C	4	–	–	–	–	–	–	–
Arabidopsis lyrata ssp. *petraea* 14	Met	3C	3	–	–	–	–	–	–	–
Arabidopsis lyrata ssp. *petraea* 15	Met	3C	4	–	–	–	–	–	–	–
Arabidopsis lyrata ssp. *petraea* 16	Met	3C	4	–	–	–	–	–	–	–
Arabidopsis lyrata ssp. *petraea* 17	Met	3C	4	–	–	–	–	–	–	–
Arabidopsis lyrata ssp. *petraea* 19	Met	3C	–	–	3	–	–	–	–	–
Arabidopsis lyrata ssp. *petraea* 18	Met	3C	4	–	–	–	–	–	–	–
Arabidopsis halleri 1	Met	5C+	–	–	–	–	–	–	–	–
Arabidopsis halleri 2	Met	5C+	–	–	–	–	–	–	–	–
Arabidopsis halleri 3	Met	5C+	–	–	–	–	–	–	–	–
Arabidopsis halleri 4	Met	5C+	–	–	–	–	–	–	–	–
Arabidopsis halleri 5	Met	5C+	–	–	–	–	–	–	–	–
Arabidopsis arenosa 1	Met	5C+	–	–	–	–	–	–	–	–
Arabidopsis arenosa 2	Met	5C+	–	–	–	–	–	–	–	–
Arabidopsis arenosa 3	Met	5C+	–	–	1	–	–	–	–	–
Arabidopsis cebennensis	Met	4C	–	–	–	–	–	–	–	4
Arabidopsis croatica	Met	3C	–	–	–	–	4	–	–	–
Boechera drummondii	Met	4C	–	–	–	1	–	2	–	–
Boechera holboellii	Val	3C	–	–	1	–	–	–	–	–
Cheesemania	Met	3C	–	3	–	–	–	–	–	–
Crucihimalaya wallichii	Met	3C	–	3	–	–	3	–	–	–
Halimolobos perplexa	Met	3C	–	3	–	–	3	–	–	–
Halimolobos whitedii	Met	3C	–	3	–	–	3	–	–	–
Olimarabidopsis pumila 1	Met	3C	4	–	–	–	–	–	–	–
Olimarabidopsis pumila 2	Met	3C	–	1	4	1	–	–	–	–
Polyctenium fremontii 1	Met	3C	–	4	1	–	–	1	–	–
Polyctenium fremontii 2	Met	3C	–	4	–	–	1	–	–	–
Turritis glabra	Met	5C+	–	–	–	–	–	–	–	–
Arabis pendula	n/a	n/a	–	–	–	–	–	–	–	–
Camelina microcarpa	n/a	n/a	–	–	–	–	–	–	–	–
Neslia paniculata	n/a	n/a	–	–	–	–	–	–	–	–
Barbarea intermedia	Phe	n/a	–	–	–	–	–	–	–	–
Barbarea verna	Phe	n/a	–	–	–	–	–	–	–	–
Barbarea vulgaris	Phe	n/a	–	–	–	–	–	–	–	–
Cardamine amara	Leu	n/a	–	–	–	–	–	–	–	–
Rorippa islandica	Trp	5C+	–	–	–	–	–	–	–	–
Lepidium campestre	Tyr[c]	4C	–	–	–	–	–	1	–	–
Lepidium perfoliatum	Met	3C	–	3	–	1	3	2	–	–
Erysimum cheiri 1	Met	3C	–	–	3	–	–	–	–	–
Erysimum cheiri 2	Met	3C	–	–	3	–	–	1	2	–
Erysimum perovskianum 1	Met	3C	–	2	3	–	–	–	–	–
Erysimum perovskianum 2	Met	3C	–	–	2	–	–	–	–	–
Erysimum repandum	Met	3C	–	3	–	1	–	–	–	–

(*continued on next page*)

Table 1 (*continued*)

Species	5MSOP	5MSOOP	4-Pent	6MSOH	6MSOOH	6MTH	7MSOH	7MSOOH	7MTH	8MSOO	8MTO
Arabidopsis thaliana[b]	–	–	+	+	–	–	+	–	+	+	+
Arabidopsis lyrata ssp. *lyrata* 1	–	–	–	–	–	–	2	–	1	–	–
Arabidopsis lyrata ssp. *lyrata* 2	–	–	–	1	–	–	3	–	1	–	–
Arabidopsis lyrata ssp. *lyrata* 3	–	–	–	–	–	–	3	–	–	3	–
Arabidopsis lyrata ssp. *lyrata* 4	–	–	–	–	–	–	1	–	–	–	–
Arabidopsis lyrata ssp. *lyrata* 5	–	–	–	–	–	–	1	–	1	1	1
Arabidopsis lyrata ssp. *lyrata* 6	–	–	–	–	–	–	1	–	1	–	–
Arabidopsis lyrata ssp. *lyrata* 7	–	–	–	–	–	–	1	–	1	1	1
Arabidopsis lyrata ssp. *lyrata* 8	–	–	–	1	–	–	2	–	2	2	–
Arabidopsis lyrata ssp. *petraea* 1	2	1	–	4	–	–	2	–	–	–	–
Arabidopsis lyrata ssp. *petraea* 2	1	–	–	4	–	–	2	–	–	–	–
Arabidopsis lyrata ssp. *petraea* 3	–	–	–	–	–	–	1	–	–	–	–
Arabidopsis lyrata ssp. *petraea* 4	–	–	–	3	–	1	1	–	–	–	–
Arabidopsis lyrata ssp. *petraea* 5	–	–	–	3	–	1	1	–	–	–	–
Arabidopsis lyrata ssp. *petraea* 6	–	–	–	1	–	–	4	–	1	1	–
Arabidopsis lyrata ssp. *petraea* 7	–	–	–	–	–	–	1	–	–	–	–
Arabidopsis lyrata ssp. *petraea* 8	–	–	–	2	–	–	2	–	–	–	–
Arabidopsis lyrata ssp. *petraea* 9	–	–	–	–	–	–	–	–	–	1	–
Arabidopsis lyrata ssp. *petraea* 10	–	–	–	1	–	–	2	–	–	–	–
Arabidopsis lyrata ssp. *petraea* 11	–	–	–	–	–	1	1	–	1	1	–
Arabidopsis lyrata ssp. *petraea* 12	–	–	–	1	–	–	2	–	1	–	–
Arabidopsis lyrata ssp. *petraea* 13	–	–	–	–	–	–	1	–	–	1	–
Arabidopsis lyrata ssp. *petraea* 14	–	–	–	–	–	–	1	–	–	1	–
Arabidopsis lyrata ssp. *petraea* 15	–	–	–	–	–	–	1	–	–	1	–
Arabidopsis lyrata ssp. *petraea* 16	–	–	–	–	–	1	2	–	1	1	–
Arabidopsis lyrata ssp. *petraea* 17	–	–	–	–	–	–	1	–	–	1	–
Arabidopsis lyrata ssp. *petraea* 19	–	–	–	–	–	–	1	–	–	4	–
Arabidopsis lyrata ssp. *petraea* 18	–	–	–	–	–	–	–	–	–	–	–
Arabidopsis halleri 1	–	–	–	4	–	2	3	–	2	–	–
Arabidopsis halleri 2	–	–	–	3	–	3	3	–	3	–	–
Arabidopsis halleri 3	–	–	–	1	–	1	4	–	3	3	1
Arabidopsis halleri 4	–	–	–	2	–	2	2	–	2	–	–
Arabidopsis halleri 5	–	–	–	4	–	–	3	–	–	2	–
Arabidopsis arenosa 1	–	–	–	–	–	–	1	–	–	4	–
Arabidopsis arenosa 2	–	–	–	–	–	–	1	–	–	4	–
Arabidopsis arenosa 3	–	–	–	–	–	–	1	–	1	4	1
Arabidopsis cebennensis	–	–	1	–	–	1	2	–	1	–	–
Arabidopsis croatica	–	–	–	–	–	–	1	–	–	2	–
Boechera drummondii	–	–	–	4	–	–	1	–	1	1	2
Boechera holboellii	–	–	–	1	–	–	–	–	–	–	–
Cheesemania	–	–	–	–	–	–	2	–	–	2	–
Crucihimalaya wallichii	–	–	–	–	–	–	2	–	–	1	–
Halimolobos perplexa	–	–	–	–	–	–	–	–	–	1	–
Halimolobos whitedii	–	–	–	–	–	–	–	–	–	1	–
Olimarabidopsis pumila 1	–	–	–	–	–	–	–	–	–	1	–
Olimarabidopsis pumila 2	–	–	–	–	–	–	–	–	–	–	–
Polyctenium fremontii 1	–	–	–	1	–	–	2	–	–	1	–
Polyctenium fremontii 2	–	–	–	–	–	–	2	–	–	1	–
Turritis glabra	–	–	–	2	4	1	2	2	–	–	–
Arabis pendula	–	–	–	–	–	–	–	–	–	–	–
Camelina microcarpa	–	–	–	–	–	–	–	–	–	–	–
Neslia paniculata	–	–	–	–	–	–	–	–	–	–	–
Barbarea intermedia	–	–	–	–	–	–	–	–	–	–	–
Barbarea verna	–	–	–	–	–	–	–	–	–	–	–
Barbarea vulgaris	–	–	–	–	–	–	–	–	–	–	–
Cardamine amara	–	–	–	–	–	–	–	–	–	–	–
Rorippa islandica	–	–	–	–	–	–	1	–	–	–	–
Lepidium campestre	2	–	–	1	–	–	–	–	–	–	–
Lepidium perfoliatum	–	–	–	–	–	–	–	–	–	–	–
Erysimum cheiri 1	–	–	–	–	–	–	–	–	–	–	–
Erysimum cheiri 2	–	–	–	–	–	–	–	–	–	–	–
Erysimum perovskianum 1	–	–	–	–	–	–	–	–	–	–	–
Erysimum perovskianum 2	–	–	–	–	–	–	–	–	–	–	–
Erysimum repandum	–	–	–	–	–	–	–	–	–	–	–

Glucosinolate quantities are ranked qualitatively within a given sample from '–' (absent) to '4' (highly abundant). Abbreviations correspond to systematic glucosinolate naming as follows: 3OHP, 3-hydroxypropyl; 3MSOP, 3-methylsulfinylpropyl; 3MSOOP, 3-methylsulfonylpropyl; 3MTP, 3-methylthiopropyl; Allyl, 2-propenyl; 4MSOB, 4-methylsulfinylbutyl; 4MSOOB, 4-methylsulfonylbutyl; 3-But, 3-butenyl; 5MSOP, 5-methylsulfinylpentyl; 5MSOOP, 5-methylsulfonylpentyl; 4-Pent, 4-pentenyl; 6MSOH, 6-methylsulfinylhexyl; 6MSOOH, 6-methylsulfonylhexyl; 6MTH, 6-methylthiohexyl; 7MSOH, 7-methylsulfinylheptyl; 7MSOOH, 7-methylsulfonylheptyl; 7MTH, 7-methylthioheptyl; 8MSOO, 8-methylsulfinyloctyl; 8MTO, 8-methylthiooctyl.

[a] Functional configuration of the *GS-Elong* locus within a given accession; based on the Met-glucosinolate profile. '3C' indicates that a single round of methionine chain-elongation predominates (MAM2-like activity); '4C' indicates two rounds of methionine chain-elongation predominate (MAM1-like activity); in accessions labeled '5C+', chain lengths longer than 3C and 4C predominate. 'n/a' indicates the complete absence of Met-glucosinolates.

[b] Glucosinolate profile for *A. thaliana* vegetative tissue. '+' indicates presence of the glucosinolate in 1 or more *A. thaliana* accessions; '–' the absence of the glucosinolate in all assayed accessions (Kliebenstein et al., 2001a; Brown et al., 2003).

[c] The tyrosine-derived glucosinolate, *p*-hydroxybenzyl (Petersen et al., 2001) is the dominant glucosinolate in this species.

[*] The amino acid precursor of the predominant class of glucosinolate within a given accession; 'n/a' indicates the complete absence of glucosinolates in leaf tissue.

Table 2
Distribution of non-methionine-derived glucosinolates in the genus *Arabidopsis* and other Brassicaceae

Species	Val*	Ile*	Leu*	Trp*			Phe*		
	1ME	1MP	2MP	I3M	4MO-I3M	1MO-I3M	Benz[b]	2PE	(*S*)2OH-2PE
Arabidopsis thaliana[a]	–	–	–	+	+	+	+	–	–
Arabidopsis lyrata ssp. *lyrata* 1	–	–	–	1	1	–	–	–	–
Arabidopsis lyrata ssp. *lyrata* 2	–	–	–	1	1	–	–	–	–
Arabidopsis lyrata ssp. *lyrata* 3	–	–	–	1	–	–	1	–	–
Arabidopsis lyrata ssp. *lyrata* 4	–	–	–	1	–	–	–	–	–
Arabidopsis lyrata ssp. *lyrata* 5	–	–	–	2	1	–	–	–	–
Arabidopsis lyrata ssp. *lyrata* 6	–	–	–	2	1	–	–	–	–
Arabidopsis lyrata ssp. *lyrata* 7	–	–	–	2	1	–	–	–	–
Arabidopsis lyrata ssp. *lyrata* 8	–	–	–	3	1	2	–	–	–
Arabidopsis lyrata ssp. *petraea* 1	–	–	–	2	1	–	–	–	–
Arabidopsis lyrata ssp. *petraea* 2	–	–	–	2	1	–	–	–	–
Arabidopsis lyrata ssp. *petraea* 3	–	–	–	2	1	–	–	–	–
Arabidopsis lyrata ssp. *petraea* 4	–	–	–	1	1	–	–	–	–
Arabidopsis lyrata ssp. *petraea* 5	–	–	–	1	1	–	–	–	–
Arabidopsis lyrata ssp. *petraea* 6	–	–	–	3	2	1	–	–	–
Arabidopsis lyrata ssp. *petraea* 7	–	–	–	1	1	1	–	–	–
Arabidopsis lyrata ssp. *petraea* 8	–	–	–	1	1	–	–	–	–
Arabidopsis lyrata ssp. *petraea* 9	–	–	–	1	–	–	–	–	–
Arabidopsis lyrata ssp. *petraea* 10	–	–	–	1	1	–	–	–	–
Arabidopsis lyrata ssp. *petraea* 11	–	–	–	1	1	–	–	–	–
Arabidopsis lyrata ssp. *petraea* 12	–	–	–	3	2	1	–	–	–
Arabidopsis lyrata ssp. *petraea* 13	–	–	–	1	1	–	–	–	–
Arabidopsis lyrata ssp. *petraea* 14	–	–	–	2	1	–	–	–	–
Arabidopsis lyrata ssp. *petraea* 15	–	–	–	1	1	–	–	–	–
Arabidopsis lyrata ssp. *petraea* 16	–	–	–	1	1	–	–	–	–
Arabidopsis lyrata ssp. *petraea* 17	–	–	–	1	1	–	–	–	–
Arabidopsis lyrata ssp. *petraea* 19	–	–	–	1	–	–	–	–	–
Arabidopsis lyrata ssp. *petraea* 18	–	–	–	1	–	–	–	–	–
Arabidopsis halleri 1	–	–	–	1	1	–	–	–	–
Arabidopsis halleri 2	–	–	–	–	1	–	–	–	–
Arabidopsis halleri 3	–	–	–	–	1	–	–	–	–
Arabidopsis halleri 4	–	–	–	–	1	–	–	–	–
Arabidopsis halleri 5	–	–	–	–	1	–	–	–	–
Arabidopsis arenosa 1	–	–	1	2	1	–	–	–	–
Arabidopsis arenosa 2	–	–	–	2	1	–	–	–	–
Arabidopsis arenosa 3	–	–	–	2	1	–	–	–	–
Arabidopsis cebennensis	–	–	–	1	1	–	–	–	–
Arabidopsis croatica	–	–	–	2	1	–	–	–	–
Boechera drummondii	2	–	1	–	1	–	–	–	–
Boechera holboellii	4	2	1	–	2	1	–	–	–
Cheesemania	–	–	–	1	2	–	–	–	–
Crucihimalaya wallichii	–	–	–	–	–	–	–	–	–
Halimolobos perplexa	–	–	–	1	1	–	–	–	–
Halimolobos whitedii	–	–	–	1	1	–	–	–	–
Olimarabidopsis pumila 1	–	–	–	1	1	–	–	–	–
Olimarabidopsis pumila 2	–	–	–	–	1	–	–	–	–
Polyctenium fremontii 1	–	–	–	–	1	–	–	–	–
Polyctenium fremontii 2	–	–	–	–	1	–	–	–	–
Turritis glabra	–	–	–	–	2	–	–	–	–
Arabis pendula	–	–	–	–	–	–	–	–	–
Camelina microcarpa	–	–	–	–	–	–	–	–	–
Neslia paniculata	–	–	–	–	–	–	–	–	–
Barbarea intermedia	–	–	–	2	–	–	–	–	3
Barbarea verna	–	–	–	1	–	–	–	4	–
Barbarea vulgaris	–	–	–	1	–	–	–	–	4
Cardamine amara	–	–	3	–	1	–	2	–	–
Rorippa islandica	–	–	–	4	–	–	–	–	–
Lepidium campestre	–	–	–	–	1	–	–	–	–
Lepidium perfoliatum	–	–	–	1	1	–	–	–	–
Erysimum cheiri 1	–	–	–	–	–	–	–	–	–
Erysimum cheiri 2	–	–	–	–	–	–	–	–	–

(*continued on next page*)

Table 2 (*continued*)

Species	Val*	Ile*	Leu*	Trp*			Phe*		
	1ME	1MP	2MP	I3M	4MO-I3M	1MO-I3M	Benz[b]	2PE	(*S*)2OH-2PE
Erysimum perovskianum 1	–	–	–	–	1	–	–	–	–
Erysimum perovskianum 2	–	–	–	–	–	–	–	–	–
Erysimum repandum	–	–	–	–	1	–	–	–	–

Glucosinolate quantities are ranked qualitatively within a given sample from '–' (absent) to '4' (highly abundant). Abbreviations correspond to systematic glucosinolate naming as follows: 1ME, 1-methylethyl; 1MP, 1-methylpropyl; 2MP, 2-methylpropyl; I3M, indol-3-yl-methyl; 4MO-I3M, 4-methoxy-indol-3-yl-methyl; 1MO-I3M, 1-methoxy-indol-3-yl-methyl; Benz, benzyl; 2PE, 2-phenylethyl; (*S*)2OH-2PE, (*S*)-2-hydroxy-2-phenylethyl.

[a] Glucosinolate profile for *A. thaliana* vegetative tissue. '+' indicates presence of the glucosinolate in 1 or more *A. thaliana* accessions; '–' the absence of the glucosinolate in all assayed accessions (Kliebenstein et al., 2001a; Brown et al., 2003).

[b] Benzyl glucosinolate is present at extremely low levels in the leaves of *A. thaliana* (Kliebenstein et al., 2001a; Reichelt et al., 2002); a score of '–' in related species may indicate that the compound is present at a level below our detection limit.

[*] The amino acid precursor of the indicated glucosinolate species; 'n/a' indicates the complete absence of glucosinolates in leaf tissue.

2.1.3. Variation in Met-glucosinolates: Side-chain modification

In *A. thaliana*, methylthioalkyl glucosinolates are oxidized to produce methylsulfinylalkyl glucosinolates (Fig. 2). This feature of glucosinolate metabolism is ubiquitous throughout the genus and in all analyzed taxa that express Met-glucosinolates in their vegetative tissues (Table 1). Further oxidation to methylsulfonyl glucosinolates (Fig. 2) has not been observed in *A. thaliana* but is detected in other members of the genus (*A. lyrata* and *A. arenosa*, Table 1) as well as in a subset of the closely related genera (Table 1). Nothing is known regarding the biochemistry or the genetics behind this oxidation step and the lack of phylogenic pattern supports suggestions that these reactions reflect lineage-specific physiological constraints.

The methylsulfinylalkyl glucosinolates of *A. thaliana* are modified by the activity of the *ALK-OHP* complex locus (Fig. 2), which encodes a series of directly repeated 2-oxoglutarate-dependent dioxygenases. In this species, different accessions display one of three phenotypic states: they lack ALK-OHP activity and produce methylsulfinylalkyl glucosinolates; display AOP2 activity, thus producing alkenyl glucosinolates; or display AOP3 activity and produce hydroxyalkyl glucosinolates (Fig. 2; Kliebenstein et al., 2001a,b). Similar patterns are observed throughout the genus *Arabidopsis*. Differences in the activity of *ALK-OHP* parallel the geographic separation of the two *A. lyrata* subspecies (ssp.) presented in Table 1. When ALK-OHP activity is present, the European subspecies, *A. lyrata* ssp. *petraea*, produces predominantly hydroxyalkyl glucosinolates (Table 1). In contrast, North American *A. lyrata* ssp. *lyrata* produces mainly alkenyl glucosinolates (Table 1). Among the other *Arabidopsis* species, *A. cebennensis* and *A. croatica* both produce alkenyl glucosinolates, while the low levels of short chain glucosinolates in *A. halleri* and *A. arenosa* prevent an assessment of ALK-OHP activity. cDNAs corresponding to *AOP1* and *AOP2* have been obtained from *A. lyrata* and an *AOP1* cDNA has been recovered from *A. halleri* (unpublished Genbank sequences- AF427857; AF418239; AF418252). The identification of *AOPx* cDNAs in *A. lyrata* and *A. halleri* and the placement of these species relative to *A. thaliana* in the phylogeny of *Arabidopsis* (Fig. 3(a)) suggest that *ALK-OHP* is present as a complex throughout the genus, and that some components of this multigene family have lost function in daughter lineages.

Outside the genus *Arabidopsis*, only five out of the 16 genera in Table 1 display ALK-OHP activity in vegetative tissue. Of these, only one species, *Olimarabidopsis pumila* produces a hydroxyalkyl glucosinolate (AOP3 activity) in vegetative tissues, while the remaining genera produce alkenyl glucosinolates (AOP2 activity). Given that in a majority of instances only one or two accessions have been assayed for glucosinolates in each genus outside of *Arabidopsis*, we cannot exclude the possibility that a more extensive survey might detect greater intra-/interspecific variation in ALK-OHP activity.

In the phylogeny of the Brassicaceae, the genus *Brassica*, which last shared a common ancestor with *A. thaliana* approximately 23 million years ago (Yang et al., 1999; Koch et al., 2001), is basal to the genera presented here. *Brassica* contains *ALK-OHP* as a complex locus (*BoALK-OHP*, Gao et al., 2004) showing that diversification of the *ALK-OHP* gene family predates the divergence of *Brassica* and *Arabidopsis*. However, analysis of the *Brassica oleracea ALK-OHP* locus demonstrates that variation in ALK-OHP activity need not be the result of polymorphisms affecting transcriptional regulation as is observed in *A. thaliana* (Kliebenstein et al., 2001b); *BoALK-OHP* lacks *AOP3* entirely, and polymorphisms affecting AOP2 activity are structural (i.e. deletions of coding sequence) rather than regulatory (Li and Quiros, 2003; Gao et al., 2004).

2.2. Divergence of non-Met-glucosinolate profiles

Non-Met-glucosinolates, derived from Trp and Phe, are detectable in the vegetative tissues of *A. thaliana*

(Kliebenstein et al., 2001a; Brown et al., 2003) and the profiles of these compounds in closely related species do not differ substantially from the pattern observed in *A. thaliana* (Table 2). However, the leaves of *A. thaliana* completely lack glucosinolates derived from the branched-chain amino acids Val, Ile, and Leu (Kliebenstein et al., 2001a; Table 2). This phenotype is maintained in the majority of related species with four exceptions: *A. arenosa*, *Boechera drummondii*, *B. holboellii*, and *Cardamine amara* (Table 2). As *A. thaliana* produces branched-chain aliphatic glucosinolates in seeds (Kliebenstein et al., 2001a), the character change identified in these species likely represents a regulatory change resulting in a novel expression pattern for branched-chain amino acid specific components of the core glucosinolate biosynthetic pathway, possibly affecting a *CYP79* or a *CYP83*. While the polymorphism associated with *B. drummondii* and *B. holboellii* appears to be a single, shared change resulting in the gain of branched-chain glucosinolates in leaves (Fig. 3(a)), the observed changes in *A. arenosa* and *C. amara* represent parallel evolution of leaf expression patterns, relative to the ancestral lack of leaf expression among most *Arabidopsis* relatives.

2.3. Geographic and historical distribution of glucosinolate metabolism in A. lyrata

Together with accessions from four additional *Arabidopsis* species, *A. lyrata* ssp. *lyrata* and *A. lyrata* ssp. *petraea* accessions were subjected to a cluster analysis using the glucosinolate phenotypic components presented in Tables 1 and 2 as characters. The resulting tree (Fig. 3(b)), which graphically depicts the biochemical similarity among the accessions, indicates that North American and European *A. lyrata* accessions can be divided into five contemporary phenotypes. Further, these phenotypes, with some notable exceptions, recapitulate the geographic origins of the plants (Fig. 3(b) and (c)). Two *A. petraea* accessions from Central Europe (Fig. 3(b) and (c); orange labels), however, are distinct from other Central European *A. petraea*; their glucosinolate metabolism is more similar to that of *A. arenosa* (Fig. 3(b)). As *A. lyrata* ssp. *petraea* and *A. arenosa* are sympatric, this observation may suggest gene-flow between the species, but few data exist to directly address this issue. Surprisingly, two North American accessions, *A. lyrata* ssp. *lyrata* 1 and 2, and one Eastern European accession, *A. lyrata* ssp. *petraea* 6, form a minor clade (Fig. 3(b), green labels). The meaning of this observation is unclear; the gross similarity in glucosinolate biochemistry may be indicative of parallel evolution resulting from comparable environmental and/or ecological pressures. Alternatively, these three accessions may represent an ancestral phenotype that existed prior to the radiation of *A. lyrata* throughout Europe and North America.

2.4. Genetic constraints and glucosinolate evolution

Comparisons within and among species provide information regarding evolutionary constraints on physiological and developmental change. Such constraints may involve trade-offs among phenotypic traits controlled by a single gene (Kroymann et al., 2003; Baucom and Mauricio, 2004; Tian et al., 2003), parallel evolution driven by orthologous loci in different species (Yoon and Baum, 2004; Schluter et al., 2004), or losses of function in biochemical pathways that preclude future evolutionary change (Zufall and Rausher, 2004; Rutter and Zufall, 2004). In the first case, while glucosinolates reliably provide defense against generalist insect herbivores, the defensive capacity of the glucosinolates is constrained by the tendency of the same compounds to attract specialist herbivores that detoxify these metabolites (Raybould and Moyes, 2001; Moyes and Raybould, 2001; Kroymann and Mitchell-Olds, 2004; Ratzka et al., 2002). Parallel evolution at orthologous loci is also found in glucosinolate biosynthesis: *AOP* genes encode polymorphic glucosinolate modification enzymes that co-segregate with glucosinolate differences in both *A. thaliana* (Kliebenstein et al., 2001b) and *B. oleracea* (Li and Quiros, 2003). Finally, the third constraint mechanism, where loss of function precludes future evolutionary change, has been clearly documented within *A. thaliana*. Naturally occurring loss-of-function mutations at *GS-Elong*, *ALK-OHP*, and *Epithiospecifier Protein* segregate among *A. thaliana* ecotypes and influence natural variation for insect resistance (Kroymann et al., 2001, 2003; Kliebenstein et al., 2001b; Lambrix et al., 2001). Although similar loss-of-function mutations have not yet been shown to constrain glucosinolate evolution for an entire species, the results of our current work suggests potential examples, such as members of the Neslia group and the Cardamine alliance, where this mechanism may be operating.

3. Experimental

3.1. Plant material

All plants were grown under standard laboratory conditions for *A. thaliana* as described previously (Kliebenstein et al., 2001a).

3.2. Glucosinolate extraction and purification

Glucosinolate extraction and purification was carried out according to Kroymann et al. (2001).

3.3. HPLC and glucosinolate identification

High-pressure liquid chromatography (HPLC) with a diode-array detector was performed as described in Kliebenstein et al. (2001a) and Reichelt et al. (2002). Preliminary identification of *desulfo*-glucosinolates was carried out by comparison of retention times and UV absorption spectra with those of known standards. Final identification of all peaks was based on an electrospray mass spectrometry (LC/MS) or LC/MS/MS as described by van Dam et al. (2004).

3.4. Statistics

Multivariate clustering of data on glucosinolate type and relative concentration was performed using UPGMA in the SYSTAT statistical package (Systat Software GmbH, Erkrath, Germany). This graphical portrayal is a summary of the patterns of biochemical similarity, and does not necessarily imply phylogenetic relationships or cladogenesis.

Acknowledgements

The authors would like to thank R. Oyama for the supplemented phylogeny of the Brassicaceae and K. Weniger, T. Kruegel, and the MPI-CE greenhouse-staff for plant growth and maintenance. We also wish to thank J. Tokuhisa, J. Bishop, M. Clauss, and C. Wheat for constructive and insightful discussions on the content of the manuscript. The authors gratefully acknowledge the DFG (grant GE 1126/1-3 to J.G.; grant KR 2237/2-1 to J.K.), the EU (grant QRLT-2000-01097 to T.M.O.), the US NSF (grant DEB-9527725 to T.M.O.), and the Max Planck Society for financial support.

References

Al-Shehbaz, I.A., Steve, L., O'Kane, J., 2002. Taxononmy and phylogeny of *Arabidopsis* (Brassicaceae). In: Somerville, C.R., Meyerowitz, E.M. (Eds.), The Arabidopsis Book. American Society of Plant Biologists, Rockville, MD, pp. 1–22.

Bak, S., Feyereisen, R., 2001. The involvement of two p450 enzymes, CYP83B1 and CYP83A1, in auxin homeostasis and glucosinolate biosynthesis. Plant Physiol. 127, 108–118.

Baucom, R.S., Mauricio, R., 2004. Fitness costs and benefits of novel herbicide tolerance in a noxious weed. Proc. Natl. Acad. Sci. USA 101, 13386–13390.

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.

Castro, A., Aires, A., Rosa, E., Bloem, E., Stulen, I., De Kok, L.J., 2004. Distribution of glucosinolates in *Brassica oleracea* cultivars. Phyton 44, 133–143.

Charron, C.S., Sams, C.E., 2004. Glucosinolate content and myrosinase activity in rapid-cycling *Brassica oleracea* grown in a controlled environment. J. Am. Soc. Hortic. Sci. 129, 321–330.

Chen, S., Glawischnig, E., Jorgensen, K., Naur, P., Jorgensen, B., Jorgensen, B., Olsen, C.-E., Hansen, C.H., Rasmussen, H., Pickett, J.A., Halkier, B.A., 2003. CYP79F1 and CYP79F2 have distinct functions in the biosynthesis of aliphatic glucosinolates in *Arabidopsis*. Plant J. 33, 923–937.

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.

Doughty, K.J., Porter, A.J.R., Morton, A.M., Kiddle, G., Bock, C.H., Wallsgrove, R., 1991. Variation in the glucosinolate content of oilseed rape *Brassica-Napus* L. leaves II. Response to infection by *Alternaria-Brassicae* Berk. Sacc. Ann. Appl. Biol. 118, 469–478.

Eriksson, S., Ek, B., Xue, J., Rask, L., Meijer, J., 2001. Identification and characterization of soluble and insoluble myrosinase isoenzymes in different organs of *Sinapis alba*. Physiol. Plantarum 111, 353–364.

Fahey, J.W., Zalcmann, A.T., Talalay, P., 2001. The chemical diversity and distribution of glucosinolates and isothiocyanates among plants. Phytochemistry 56, 5–51.

Field, B., Cardon, G., Traka, M., Botterman, J., Vancanneyt, G., Mithen, R., 2004. Glucosinolate and amino acid biosynthesis in *Arabidopsis*. Plant Physiol. 135, 828–839.

Font, R., Del Rio, M., Fernandez-Martinez, J.M., de Haro-Bailon, A., 2004. Use of near-infrared spectroscopy for screening the individual and total glucosinolate contents in Indian mustard seed (*Brassica juncea* L. Czern. & Coss.). J. Agric. Food Chem. 52, 3563–3569.

Gao, M., Li, G., Yang, B., McCombie, W.R., Quiros, C.F., 2004. Comparative analysis of a *Brassica* BAC clone containing several major aliphatic glucosinolate genes with its corresponding *Arabidopsis* sequence. Genome 47, 666–679.

Geshi, N., Andreasson, E., Meijer, J., Rask, L., Brandt, A., 1998. Co-localization of myrosinase- and myrosinase-binding proteins in grains of myrosin cells in cotyledon of *Brassica napus* seedlings. Plant Physiol. Biochem. 36, 583–590.

Graser, G., Oldham, N.J., Brown, P.D., Temp, U., Gershenzon, J., 2001. The biosynthesis of benzoic acid glucosinolates esters in *Arabidopsis thaliana*. Phytochemistry 57, 23–32.

Hansen, C.H., Wittstock, U., Olsen, C.E., Hick, A.J., Pickett, J.A., Halkier, B.A., 2001. Cytochrome P450 CYP79F1 from *Arabidopsis* catalyzes the conversion of dihomomethionine and trihomomethionine to the corresponding aldoximes in the biosynthesis of aliphatic glucosinolates. J. Biol. Chem. 14, 11078–11085.

Hemm, M.R., Ruegger, M.O., Chapple, C., 2003. The *Arabidopsis ref2* mutant is defective in the gene encoding CYP83A1 and shows both phenylpropanoid and glucosinolate phenotypes. Plant Cell 15, 179–194.

Hull, A.K., Vij, R., Celenza, J.L., 2000. *Arabidopsis* cytochrome P450s that catalyze the first step of tryptophan-dependent indole-3-acetic acid biosynthesis. Proc. Natl. Acad. Sci. USA 97, 2379–2384.

Husebye, H., Chadchawan, S., Winge, P., Thangstad, O.P., Bones, A.M., 2002. Guard cell- and phloem idioblast-specific expression of thioglucoside glucohydrolase 1 (myrosinase) in *Arabidopsis*. Plant Physiol. 128, 1180–1188.

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

Kjaer, A., Schuster, A., 1972. Glucosinolates in seeds of *Neslia-Paniculata*. Phytochemistry 11, 3045–3048.

Kliebenstein, D.J., Kroymann, J., Brown, P., Figuth, A., Pedersen, D., Gershenzon, J., Mitchell-Olds, T., 2001a. Genetic control of

natural variation in *Arabidopsis* glucosinolate accumulation. Plant Physiol. 126, 811–825.

Kliebenstein, D.J., Lambrix, V.M., Reichelt, M., Gershenzon, J., Mitchell-Olds, T., 2001b. Gene duplication in the diversification of secondary metabolism: tandem 2-oxoglutarate-dependent dioxygenases control glucosinolate biosynthesis in *Arabidopsis*. Plant Cell 13, 681–693.

Koch, M., Haubold, B., Mitchell-Olds, T., 2001. Molecular systematics of the Brassicaceae: evidence from coding plastidic matK and nuclear Chs sequences. Am. J. Bot. 88, 534–544.

Kroymann, J., Mitchell-Olds, T., 2004. Function and evolution of an *Arabidopsis* insect resistance QTL. In: Tikhonovich, I., Lugtenberg, B., Provorov, N. (Eds.), Biology of Molecular Plant–Microbe Interactions. APS Press, Saint Paul, MN, pp. 259–262.

Kroymann, J., Textor, S., Tokuhisa, J.G., Falk, K.L., Bartram, S., Gershenzon, J., Mitchell-Olds, T., 2001. A gene controlling variation in *Arabidopsis* glucosinolate composition is part of the methionine chain elongation pathway. Plant Physiol. 127, 1077–1088.

Kroymann, J., Donnerhacke, S., Schnabelrauch, D., Mitchell-Olds, T., 2003. Evolutionary dynamics of an *Arabidopsis* insect resistance quantitative trait locus. Proc. Natl. Acad. Sci. USA 100, 14587–14592.

Lambrix, V., Reichelt, M., Mitchell-Olds, T., Kliebenstein, D.J., Gershenzon, J., 2001. The *Arabidopsis* epithiospecifier protein promotes the hydrolysis of glucosinolates to nitriles and influences *Trichoplusia ni* herbivory. Plant Cell 13, 2793–2807.

Li, G., Quiros, C.F., 2003. In planta side-chain glucosinolate modification in *Arabidopsis* by introduction of dioxygenase *Brassica* homolog BoGSL-ALK. Theor. Appl. Genet. 106, 1116–1121.

Louda, S.M., Rodman, J.E., 1983. Concentration of glucosinolates in relation to habitat and insect herbivory for the native crucifer *Cardamine-Cordifolia*. Biochem. Syst. Ecol. 11, 199–208.

Marazzi, C., Staedler, E., 2004. *Arabidopsis thaliana* leaf-surface extracts are detected by the cabbage root fly (*Delia radicum*) and stimulate oviposition. Physiol. Entomol. 29, 192–198.

Mikkelsen, M.D., Hansen, C.H., Wittstock, U., Halkier, B.A., 2000. Cytochrome P450 CYP79B2 from *Arabidopsis* catalyzes the conversion of tryptophan to indole-3-acetaldoxime, a precursor of indole glucosinolates and indole-3-acetic acid. J. Biol. Chem. 275, 33712–33717.

Mikkelsen, M.D., Petersen, B.L., Olsen, C.E., Halkier, B.A., 2002. Biosynthesis and metabolic engineering of glucosinolates. Amino Acids 22, 279–295.

Mikkelsen, M.D., Petersen, B.L., Glawischnig, E., Jensen, A.B., Andreasson, E., Halkier, B.A., 2003. Modulation of CYP79 genes and glucosinolate profiles in *Arabidopsis* by defense signaling pathways. Plant Physiol. 131, 298–308.

Mikkelsen, M.D., Naur, P., Halkier, B.A., 2004. *Arabidopsis* mutants in the C-S lyase of glucosinolate biosynthesis establish a critical role for indole-3-acetaldoxime in auxin homeostasis. Plant J. 37, 770–777.

Mitchell-Olds, T., Al-Shehbaz, I.A., Koch, M., Sharbel, T., 2004. Crucifer evolution in the post-genomic era. In: Henry, R. (Ed.), Diversity and Evolution of Plants – Genotype and Phenotype Variation in Higher Plants. CABI Press, Wallingford, UK, pp. 119–138.

Moyes, C.L., Raybould, A.F., 2001. The role of spatial scale and intraspecific variation in secondary chemistry in host–plant location by *Ceutorhynchus assimilis* (Coleoptera: Curculionidae). P. Roy. Soc. Lond. B Biol. 268, 1567–1573.

Naur, P., Petersen, B.L., Mikkelsen, M.D., Bak, S., Rasmussen, H., Olsen, C.-E., Halkier, B.A., 2003. CYP83A1 and CYP83B1, two nonredundant cytochrome P450 enzymes metabolizing oximes in the biosynthesis of glucosinolates in *Arabidopsis*. Plant Physiol. 133, 63–72.

Petersen, B.L., Andreasson, E., Bak, S., Agerbirk, N., Halkier, B.A., 2001. Characterization of transgenic *Arabidopsis thaliana* with metabolically engineered high levels of *p*-hydroxybenzylglucosinolate. Planta 212, 612–618.

Piotrowski, M., Schemenewitz, A., Lopukhina, A., Muller, A., Janowitz, T., Weil, E.W., Oecking, C., 2004. Desulfoglucosinolate sulfotransferases from *Arabidopsis thaliana* catalyze the final step in the biosynthesis of the glucosinolate core structure. J. Biol. Chem. 279, 50717–50725.

Pivnick, K.A., Jarvis, B.J., Slater, G.P., 1994. Identification of olfactory cues used in host–plant finding by diamondback moth, *Plutella xylostella* (Lepidoptera: Plutellidae). J. Chem. Ecol. 20, 1407–1427.

Quiros, C.F., Grellet, F., Sadowski, J., Suzuki, T., Li, G., Wroblewski, T., 2001. *Arabidopsis* and *Brassica* comparative genomics: sequence, structure and gene content in the ABI1-Rps2-Ck1 chromosomal segment and related regions. Genetics 157, 1321–1330.

Ratzka, A., Vogel, H., Kliebenstein, D.J., Mitchell-Olds, T., Kroymann, J., 2002. Disarming the mustard oil bomb. Proc. Natl. Acad. Sci. USA 99, 11223–11228.

Raybould, A.F., Moyes, C.L., 2001. The ecological genetics of aliphatic glucosinolates. Heredity 87, 383–391.

Reichelt, M., Brown, P.D., Schneider, B., Oldham, N.J., Stauber, E., Tokuhisa, J., Kliebenstein, D.J., Mitchell-Olds, Gershenzon, J., 2002. Benzoic acid glucosinolate esters and other glucosinolates from *Arabidopsis thaliana*. Phytochemistry 59, 663–671.

Reintanz, B., Lehnen, M., Reichelt, M., Gershenzon, J., Kowalczyk, M., Sandberg, G., Godde, M., Uhl, R., Palme, K., 2001. bus, A bushy *Arabidopsis CYP79F1* knockout mutant with abolished synthesis of short-chain aliphatic glucosinolates. Plant Cell 13, 351–367.

Rodman, J.E., Karol, K.G., Price, R.A., Sytsma, K.J., 1996. Molecules, morphology, and Dahlgren's expanded order Capparales. Syst. Bot. 21, 289–307.

Rodman, J.E., Soltis, P.S., Soltis, D.E., Sytsma, K.J., Karol, K.G., 1998. Parallel evolution of glucosinolate biosynthesis inferred from congruent nuclear and plastid gene phylogenies. Am. J. Bot. 85, 997–1006.

Rutter, M.T., Zufall, R.A., 2004. Pathway length and evolutionary constraint in amino acid biosynthesis. J. Mol. Evol. 58, 218–224.

Schluter, D., Clifford, E.A., Nemethy, M., McKinnon, J.S., 2004. Parallel evolution and inheritance of quantitative traits. Am. Nat. 163, 809–822.

Textor, S., Bartram, S., Kroymann, J., Falk, K.L., Hick, A., Pickett, J., Gershenzon, J., 2004. Biosynthesis of methionine-derived glucosinolates in *Arabidopsis thaliana*: recombinant expression and characterization of methylthioalkylmalate synthase, the condensing enzyme of the chain-elongation cycle. Planta 218, 1026–1035.

Thangstad, O.P., Gilde, B., Chadchawan, S., Seem, M., Husebye, H., Bradley, D., Bones, A.M., 2004. Cell specific, cross-species expression of myrosinases in *Brassica napus*, *Arabidopsis thaliana* and *Nicotiana tabacum*. Plant Mol. Biol. 54, 597–611.

Tian, D., Traw, M.B., Chen, J.Q., Kreitman, M., Bergelson, J., 2003. Fitness costs of R-gene-mediated resistance in *Arabidopsis thaliana*. Nature 423, 74–77.

Tokuhisa, J., de Kraker, J.-W., Textor, S., Gershenzon, J., 2004. The biochemical and molecular origins of aliphatic glucosinolate diversity in *Arabidopsis thaliana*. In: Romeo, J.T. (Ed.), Secondary Metabolism in Model Systems. Pergamon, New York, NY, pp. 19–38.

van Dam, N.M., Witjes, L., Svatos, A., 2004. Interactions between aboveground and belowground induction of glucosinolates in two wild *Brassica* species. New Phytol. 161, 801–810.

Wittstock, U., Halkier, B.A., 2000. Cytochrome P450 CYP79A2 from *Arabidopsis thaliana* L. catalyzes the conversion of L-phenylalanine

to phenylacetaldoxime in the biosynthesis of benzylglucosinolate. J. Biol. Chem. 275, 14659–14666.

Yang, Y.-W., Lai, K.-N., Tai, P.-Y., Li, W.-H., 1999. Rates of nucleotide substitution in angiosperm mitochondrial DNA sequences and dates of divergence between *Brassica* and other angiosperm lineages. J. Mol. Evol. 48, 597–604.

Yoon, H.-S., Baum, D.A., 2004. Transgenic study of parallelism in plant morphological evolution. Proc. Natl. Acad. Sci. USA 101, 6524–6529.

Zufall, R.A., Rausher, M.D., 2004. Genetic changes associated with floral adaptation restrict future evolutionary potential. Nature 428, 847–850.