

The Genetic Basis of Constitutive and Herbivore-Induced ESP-Independent Nitrile Formation in Arabidopsis^{1[W][OA]}

Meike Burow², Anja Losansky, René Müller, Antje Plock, Daniel J. Kliebenstein, and Ute Wittstock*

Institut für Pharmazeutische Biologie, Technische Universität Braunschweig, 38106 Braunschweig, Germany (M.B., A.L., R.M., A.P., U.W.); and Department of Plant Sciences, University of California, Davis, California 95616 (D.J.K.)

Glucosinolates are a group of thioglucosides that are components of an activated chemical defense found in the Brassicales. Plant tissue damage results in hydrolysis of glucosinolates by endogenous thioglucosidases known as myrosinases. Spontaneous rearrangement of the aglucone yields reactive isothiocyanates that are toxic to many organisms. In the presence of specifier proteins, alternative products, namely epithionitriles, simple nitriles, and thiocyanates with different biological activities, are formed at the expense of isothiocyanates. Recently, simple nitriles were recognized to serve distinct functions in plant-insect interactions. Here, we show that simple nitrile formation in Arabidopsis (*Arabidopsis thaliana*) ecotype Columbia-0 rosette leaves increases in response to herbivory and that this increase is independent of the known epithiospecifier protein (ESP). We combined phylogenetic analysis, a screen of Arabidopsis mutants, recombinant protein characterization, and expression quantitative trait locus mapping to identify a gene encoding a nitrile-specifier protein (NSP) responsible for constitutive and herbivore-induced simple nitrile formation in Columbia-0 rosette leaves. AtNSP1 is one of five Arabidopsis ESP homologues that promote simple nitrile, but not epithionitrile or thiocyanate, formation. Four of these homologues possess one or two lectin-like jacalin domains, which share a common ancestry with the jacalin domains of the putative Arabidopsis myrosinase-binding proteins MBP1 and MBP2. A sixth ESP homologue lacked specifier activity and likely represents the ancestor of the gene family with a different biochemical function. By illuminating the genetic and biochemical bases of simple nitrile formation, our study provides new insights into the evolution of metabolic diversity in a complex plant defense system.

Plants produce a multitude of secondary metabolites to survive in a complex biotic environment. These compounds function as toxins, repellents, and deterrents that defend plants against herbivores, antibiotics that are directed against pathogen attack, volatiles mediating tritrophic interactions, and signals that affect neighboring plants (Kessler and Baldwin, 2002; Wittstock and Gershenson, 2002; Kost and Heil, 2006; Pieterse and Dicke, 2007). Important secondary metabolites produced in the model plant Arabidopsis (*Arabidopsis thaliana*; Brassicaceae) are the glucosinolates, a group of amino acid-derived thioglucosides that are components of an activated chemical defense

system (Wittstock and Halkier, 2002; Kliebenstein, 2004; Kliebenstein et al., 2005; Fig. 1). This system is common to all plants of the Brassicales order and is involved in a variety of plant-insect and plant-pathogen interactions (Chew, 1988; Rask et al., 2000; Fahey et al., 2001; Halkier and Gershenson, 2006). Activation of the system occurs upon tissue disruption, which releases glucosinolates and their hydrolytic enzymes, myrosinases (EC 3.2.1.147), from separate storage compartments (Andréasson and Jørgensen, 2003). The hydrolysis products formed after myrosinase-catalyzed cleavage of the thioglucosidic bond of the glucosinolates affect a variety of organisms and are responsible for the different flavors of glucosinolate-containing vegetables and spices like cabbage (*Brassica oleracea* var. *capitata*), broccoli (*B. oleracea* var. *italica*), horseradish (*Aморacia rusticana*), mustard (*Sinapis alba*), and rucola (*Eruca sativa*; Chew, 1988; Rask et al., 2000; Fahey et al., 2001; Bennett et al., 2006).

The glucosinolate-myrosinase system can generate chemical diversity using independent variation in both glucosinolate biosynthesis and glucosinolate hydrolysis (Kliebenstein et al., 2005; Benderoth et al., 2006). Variation in glucosinolate biosynthesis leads to the production of more than 120 different glucosinolates from only a few amino acids (Halkier and Gershenson, 2006). This variation is amplified during glucosinolate hydrolysis, as a single glucosinolate can be hydrolyzed to different products with diverse physicochemical and biological properties. The outcome of glucosinolate

¹ This work was supported by the German Research Foundation, Priority Program 1152, Evolution of Metabolic Diversity (grant no. WI2668/2-1 to U.W.), and the National Science Foundation (grant nos. DBI 0642481 and MCB 0323759 to D.J.K.).

² Present address: Plant Biochemistry Laboratory, Department of Plant Biology and Biotechnology, and VKR Research Center for Pro-Active Plants, Faculty of Life Sciences, Copenhagen University, Thorvaldsensvej 40, DK-1871 Frederiksberg C, Denmark.

* Corresponding author; e-mail u.wittstock@tu-bs.de.

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantphysiol.org) is: Ute Wittstock (u.wittstock@tu-bs.de).

^[W] The online version of this article contains Web-only data.

^[OA] Open Access articles can be viewed online without a subscription.

www.plantphysiol.org/cgi/doi/10.1104/pp.108.130732

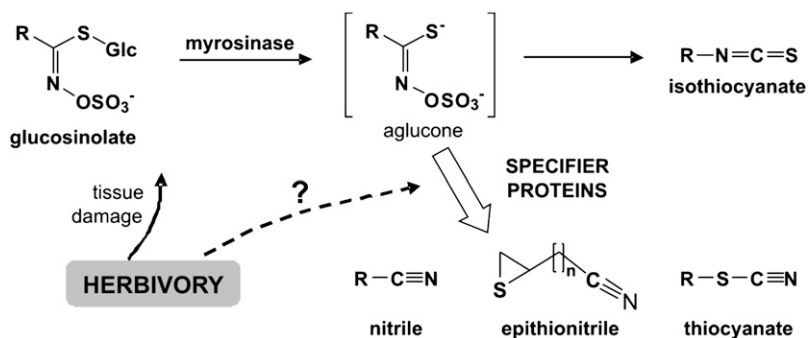


Figure 1. Glucosinolate hydrolysis. Tissue damage as caused by chewing herbivores results in myrosinase-catalyzed hydrolysis of glucosinolates, yielding Glc and unstable aglucone. Whereas the subsequent formation of simple nitriles, epithionitriles, and/or organic thiocyanates requires the action of specifier proteins, spontaneous rearrangement of the aglucone intermediates leads to isothiocyanates. The different types of hydrolysis products differ in their biological activities. Some products have been shown to act as chemical defenses against insect herbivores, but it had remained an open question whether herbivory affects the outcome of glucosinolate hydrolysis.

hydrolysis depends on the structure of the glucosinolate side chain and the presence of supplementary proteins (Tookey, 1973; Chew, 1988; Lambrix et al., 2001; Wittstock et al., 2003; Zhang et al., 2006; Burow et al., 2007a; Fig. 1). The most intensely studied hydrolysis products, the isothiocyanates, are formed through a spontaneous rearrangement of the glucosinolate aglucone in the absence of supplementary proteins. Glucosinolate-derived isothiocyanates are very reactive and have been shown to be toxic to microorganisms, nematodes, and insects (Lichtenstein et al., 1962; Tierens et al., 2001; Agrawal and Kurashige, 2003; Lazzeri et al., 2004; for review, see Wittstock et al., 2003). They have also attracted a lot of interest as cancer-preventive components of the human diet (Fahey et al., 1997; Van Poppel et al., 2000).

Plant and insect proteins known as specifier proteins promote the rearrangement of the aglucone released by myrosinase to products other than isothiocyanates, namely to simple nitriles, epithionitriles, and organic thiocyanates, without having hydrolytic activity on glucosinolates themselves (Tookey, 1973; for review, see Wittstock and Burow, 2007; Fig. 1). The full-length amino acid sequences of only three plant specifier proteins are known, namely those of the epithiospecifier proteins (ESPs) from *Arabidopsis* (Lambrix et al., 2001) and *Brassica oleracea* (Matusheski et al., 2006) and of the thiocyanate-forming protein (TFP) from *Lepidium sativum* (Burow et al., 2007a). These proteins share 63% to 67% amino acid sequence identity and form a small protein family. As a structural feature, the proteins are predicted to contain a series of β -sheets representing Kelch domains typically involved in protein-protein interactions (Adams et al., 2000; Burow et al., 2006a, 2007a; Matusheski et al., 2006). The *Arabidopsis* ESP is the best-studied member of the family to date (Lambrix et al., 2001; Zabala et al., 2005; Burow et al., 2006a, 2007b, 2008; Miao and Zentgraf, 2007). In the presence of ESP, myrosinase-catalyzed hydrolysis of alkenylglucosinolates containing a terminal double bond yields

epithionitriles, which are nitriles with a terminal thiirane ring (Fig. 1). In addition, ESP promotes the formation of simple nitriles from all other glucosinolates (Lambrix et al., 2001; Burow et al., 2006a, 2008; Matusheski et al., 2006). Feeding tests show that simple nitrile production does not function as a direct defense against generalist lepidopteran herbivores, because these herbivores prefer plants with simple nitriles over those with isothiocyanates and complete development faster on nitrile-producing than on isothiocyanate-producing plants (Lambrix et al., 2001; Burow et al., 2006b; Zhang et al., 2006). Recent studies with *Pieris rapae* (Lepidoptera) suggest that simple nitriles are involved in direct and indirect defense responses against this specialist herbivore (De Vos et al., 2008; Mumm et al., 2008). The biological function of epithionitriles is unknown (Wittstock et al., 2003). Given that the function of an individual glucosinolate in plant defense is to a large extent determined by the outcome of its hydrolysis, a better knowledge of the protein machinery involved in glucosinolate hydrolysis will improve our understanding of the roles of the glucosinolate-myrosinase system in plant defense responses.

Allelic variation at the *ESP* locus allows accessions of *Arabidopsis* to differ not only in their glucosinolate profiles but also in the types of hydrolysis products formed upon tissue disruption (Lambrix et al., 2001). For example, in plants of the Landsberg *erecta* (*Ler*) accession, the *ESP* gene is transcribed, functional ESP is produced, and macerated leaves produce predominantly simple nitriles (Lambrix et al., 2001). Despite identical coding sequences, the level of *ESP* transcript in rosette leaves of the Columbia-0 (*Col-0*) accession is below 1% of that in *Ler* (Burow et al., 2007b). *ESP* activity, measured as epithionitrile formation from allylglucosinolate added to crude protein extracts, is not detectable in any *Col-0* tissue (Burow et al., 2007b). Interestingly, seedlings, rosette leaves, and roots of *Col-0* plants form simple nitriles from both exogenous allylglucosinolate and endogenous 4-methylsulfinyl-

butylglucosinolate, and this simple nitrile-forming activity is developmentally and organ-specifically regulated (Wentzell and Kliebenstein, 2008). Moreover, T-DNA insertions in the Col-0 *ESP* and *EPITHIOSPECIFIER-MODIFIER1 (ESM1)* genes do not impair simple nitrile formation in seedlings and roots (Wentzell and Kliebenstein, 2008). Taken together, this suggests the presence of an additional specifier protein responsible for simple nitrile formation. While a protein that promotes simple nitrile but not epithionitrile formation (nitrile-specifier protein [NSP]) has been identified in larvae of the specialist herbivore *P. rapae* as the major biochemical adaptation that allows the larvae to feed on glucosinolate-containing plants (Wittstock et al., 2004; Wheat et al., 2007), there is no Arabidopsis gene encoding a protein with any significant structural similarity to the *P. rapae* NSP (Burow et al., 2006a). Thus, a novel simple nitrile-forming activity is present in Arabidopsis whose genetic and biochemical basis remains to be established.

Here, we report that in addition to basal levels of nitrile-forming activity, simple nitrile formation in Arabidopsis Col-0 rosette leaves is induced by herbivory and that this herbivore-induced simple nitrile formation is independent of ESP. Using phylogenetic analysis in conjunction with a screen of Arabidopsis T-DNA mutants, recombinant protein characterization, and expression quantitative trait locus (QTL) mapping, we identify the genetic and biochemical basis of constitutive and herbivore-induced simple nitrile formation in rosette leaves. We provide evidence that a gene whose predicted product is annotated as a myrosinase-binding protein (MBP)-like protein encodes a protein that promotes simple nitrile, but not epithionitrile or thiocyanate, formation. This gene is one representative of five related Arabidopsis genes that encode specifier proteins with NSP activity.

RESULTS

Simple Nitrile Formation Is Induced upon Herbivory in Arabidopsis Col-0 Rosette Leaves

As simple nitriles serve as volatile signals in direct and indirect defense responses against *P. rapae* (De Vos et al., 2008; Mumm et al., 2008), we tested if herbivory by *P. rapae* larvae can induce simple nitrile formation in Arabidopsis Col-0. Upon feeding on Col-0 plants for 24 h, the proportion of nitriles formed from the endogenous glucosinolates in homogenates of the damaged leaves increased dramatically. This increase was most pronounced in the damaged leaves, while undamaged leaves of the same plant were only slightly but not statistically significantly affected (Fig. 2). For the predominant glucosinolate in Col-0 rosette leaves, 4-methylsulfinylbutylglucosinolate, the proportion of nitrile formed in homogenates of the herbivore-damaged leaves increased to twice the value formed in leaf homogenates of untreated plants (Fig. 2). This

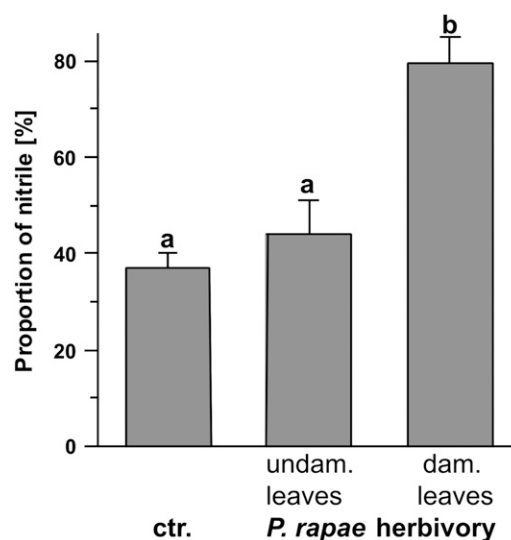


Figure 2. Herbivore-induced simple nitrile formation in Arabidopsis Col-0 rosette leaves. Glucosinolate hydrolysis products were measured in rosette leaf homogenates of 6-week-old plants that had been infested with *P. rapae*. After 24 h of herbivory, all rosette leaves were harvested. Shown is the percentage of nitrile formed from the major leaf glucosinolate 4-methylsulfinylbutylglucosinolate in relation to the total amount of hydrolysis products (nanomoles) formed from this glucosinolate. Means \pm SD from results obtained in two independent experiments ($n = 6$ per treatment class) are shown. Means were tested for significant differences by one-way ANOVA with a Tukey test (all pairwise comparisons). Means marked with a versus b were significantly different at $P < 0.001$, while the two means marked with a were not significantly different at $P > 0.5$. ctr., Control (no herbivory); dam., damaged; undam., undamaged.

change in the outcome of glucosinolate hydrolysis was not accompanied by an induction of the *ESP* transcript (Fig. 3). Moreover, the lack of epithionitrile formation from exogenous allylglucosinolate added to homogenates of herbivore-damaged leaves showed the absence of *ESP* activity (data not shown). Consequently, there must be an additional factor responsible for herbivore-induced simple nitrile formation. This factor could be a protein with structural similarity to *ESP* or a completely different protein. Alternatively, a thioglucosidase with a different product specificity than the known Arabidopsis myrosinases could be responsible for the hydrolysis of glucosinolates to simple nitriles.

Sequence Analysis of Candidate Genes Encoding Proteins with Similarity to Arabidopsis *ESP*

Based on the assumption that *ESP*-independent nitrile formation is due to a protein with similarity to *ESP*, we looked for candidate genes in the Arabidopsis genome. A BLASTN search using the Arabidopsis *Ler ESP* cDNA as a query identified six Arabidopsis genes encoding Kelch proteins with 50% to 60% amino acid sequence identity with *ESP* (Table I). Analysis of the predicted protein structures revealed that the proteins

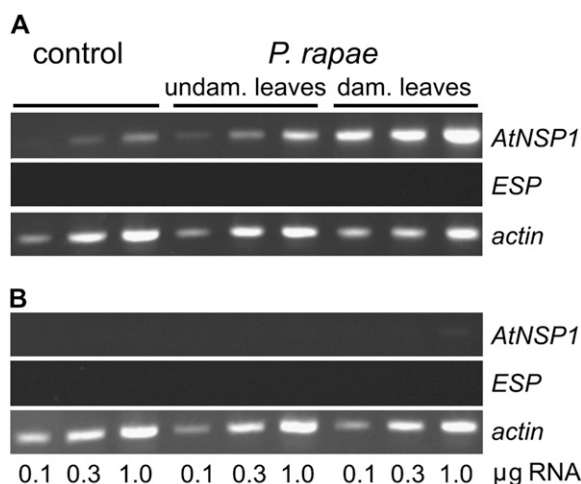


Figure 3. Expression analysis of specifier proteins by RT-PCR. Total RNA was isolated from rosette leaves of Arabidopsis Col-0 wild type (A) or SALK_072600 (B) that had been infested with *P. rapae* (control, no herbivore treatment). After 24 h of herbivory, all rosette leaves were harvested. Transcript levels of *AtNSP1* and *ESP* were analyzed by RT-PCR using actin as a control for equal amounts of RNA. RT reactions were performed on 0.1, 0.3, and 1.0 µg of RNA. The experiment was conducted on two independent sets of plants with similar results. dam., Damaged; undam., undamaged.

are composed of four or five Kelch domains (Fig. 4). Kelch domains 3 and 4 of these proteins are up to nine amino acids shorter than the respective domains in ESP. In addition, four of the six predicted candidate proteins contain one or two lectin-related jacalin domains at their N terminus with 50% to 52% identity with putative Arabidopsis MBPs (*At1g52040* and *At1g52030*; Fig. 4). Despite their annotation as MBP-like proteins, these four proteins appear to be chimeras composed of MBP- and ESP-like domains.

Arabidopsis *At3g16400* Controls Constitutive and Herbivore-Induced Simple Nitrile Formation in Rosette Leaves

To test if one of the six candidate genes was responsible for simple nitrile formation in Col-0 rosette leaves,

we performed a systematic phytochemical and biochemical screening of publicly available T-DNA insertion lines and TILLING mutants of these six genes (Supplemental Table S1). Of these mutants, a line with a T-DNA insertion in the second exon of *At3g16400* (SALK_072600) produced no or only trace amounts of simple nitriles of the endogenous glucosinolates in rosette leaf homogenates (Fig. 5). No significant changes in simple nitrile formation were detected in rosette leaf homogenates of mutant lines of the other five genes. While nitrile formation in Col-0 wild-type plants was herbivore inducible, herbivory by *P. rapae* larvae did not increase the proportion of nitriles formed upon leaf homogenization in SALK_072600 plants (Fig. 5). Reverse transcription (RT)-PCR detected *At3g16400* transcript in rosette leaves of Col-0 wild-type plants but not in the SALK_072600 T-DNA mutant (Fig. 3). While feeding by *P. rapae* larvae on wild-type plants led to increased *At3g16400* transcript levels in the damaged leaves, *At3g16400* transcript levels were undetectable in the mutant line even upon herbivory (Fig. 3). Intact glucosinolate accumulation and myrosinase activity were unaltered in the SALK_072600 mutant line compared with Col-0 wild-type plants (Supplemental Table S2; Supplemental Fig. S1). Taken together, the SALK_072600 mutant data suggest that *At3g16400* encodes a protein involved in constitutive and herbivore-induced simple nitrile formation in Arabidopsis rosette leaves.

***At3g16400* Encodes a Nitrile-Specifier Protein (*AtNSP1*)**

To confirm that *At3g16400* encodes a nitrile-specifier protein, the full-length *At3g16400* cDNA was cloned from an Arabidopsis Biological Resource Center (ABRC) cDNA clone and expressed in *Escherichia coli* with an N-terminal Strep tag. The protein purified from bacterial extracts promoted the formation of the simple nitrile, but not the epithionitrile, upon myrosinase-catalyzed hydrolysis of allylglucosinolate (Fig. 6). Therefore, we designated the protein encoded by *At3g16400* *AtNSP1*. Further biochemical characterization of the purified Strep-tagged protein showed that *AtNSP1* also converts 4-methylsulfinylbutylglucosinolate, 4-methylthiobutylglucosinolate, and benzylglucosinolate

Table I. Sequence identities of *ESP* and MBP-like proteins from Arabidopsis Col-0

Full protein sequences (*At1g54040*, *At5g48180*, and *At3g07720*) and partial protein sequences corresponding to the Kelch domain region (*At3g16400*, amino acids 145–470; *At2g33070*, 146–471; *At3g16390*, 145–467; *At3g16410*, 294–619) were aligned by ClustalW using MegAlign 7.0 (DNASTAR). Amino acid sequence identities are given in percentages.

| Protein Name | <i>At1g54040</i> | <i>At3g16400</i> | <i>At2g33070</i> | <i>At3g16390</i> | <i>At3g16410</i> | <i>At5g48180</i> |
|---------------------------------|------------------|------------------|------------------|------------------|------------------|------------------|
| <i>At1g54040</i> (<i>ESP</i>) | | | | | | |
| <i>At3g16400</i> | 59.1 | | | | | |
| <i>At2g33070</i> | 59.1 | 84.7 | | | | |
| <i>At3g16390</i> | 58.7 | 91.0 | 82.0 | | | |
| <i>At3g16410</i> | 57.3 | 94.5 | 82.2 | 86.7 | | |
| <i>At5g48180</i> | 50.8 | 53.6 | 52.3 | 51.6 | 52.3 | |
| <i>At3g07720</i> | 50.0 | 52.3 | 53.2 | 50.6 | 50.5 | 55.1 |

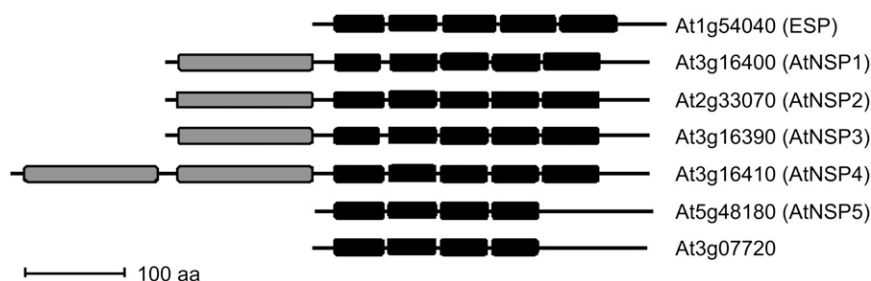


Figure 4. Predicted protein structure of ESP from Arabidopsis Ler and selected MBP-like proteins from Arabidopsis Col-0. Kelch domains (black boxes) and jacalin-like lectin domains (gray boxes) of the gene products of *At1g54040* (ESP; GenBank accession no. NP_175806), *At3g16400* (NP_566546), *At2g33070* (NP_180866), *At3g16390* (NP_566545), *At3g16410* (NP_188262), *At5g48180* (NP_568692), and *At3g07720* (NM_556316) are presented as predicted by InterProScan (<http://www.ebi.ac.uk/InterProScan/>). aa, Amino acids.

to their corresponding simple nitriles in the presence of myrosinase (Supplemental Fig. S2). AtNSP1 did not promote thiocyanate or epithionitrile formation from any of the tested substrates.

As epithionitrile formation by ESP has been demonstrated to depend on the presence of iron in vitro, we tested the iron dependence of simple nitrile formation by AtNSP1 (Fig. 7; Supplemental Fig. S3) and checked if the addition of iron would lead to detectable epithionitrile formation. Addition of 0.01 to 0.5 mM Fe²⁺ increased the proportion of nitriles formed in NSP assays performed with allyl-, 4-methylthiobutyl-, 4-methylsulfanylbutyl-, and benzylglucosinolate. However, Fe²⁺ also increased the proportion of nitriles formed in control assays that did not contain AtNSP1. Fe³⁺ had no effect on the proportion of simple nitriles formed from allyl-, benzyl-, and 4-methylthiobutylglucosinolate in our assays containing AtNSP1. In respective control assays lacking specifier proteins, Fe³⁺ does not promote simple nitrile formation (Burow et al., 2006a). To test if iron is required for the activity of AtNSP1, we added EDTA as a chelator of both Fe²⁺ and Fe³⁺ to our NSP assays. EDTA in concentrations up to 10 mM did not influence AtNSP1 activity in assays conducted with 4-methylsulfanylbutyl- and allylglucosinolate and only caused a slight, but significant, reduction in simple nitrile formation from 4-methylthiobutylglucosinolate and benzylglucosinolate (Fig. 7; Supplemental Fig. S2). Thus, simple nitrile formation by AtNSP1 is not dependent on the presence of Fe²⁺ or Fe³⁺ ion. Neither thiocyanate nor epithionitrile formation was detected in any of the reactions, further supporting the notion that AtNSP1 is specific for simple nitrile formation.

Four Homologues of AtNSP1 Have NSP, But Not ESP and TFP, Activity

To test if the other five genes identified as candidate genes also encode specifier proteins, we expressed them in *E. coli* and tested for specifier activity using allylglucosinolate and benzylglucosinolate as substrates for the myrosinase reaction. For *At3g16410*,

At3g16390, *At2g33070*, and *At5g48180*, we detected NSP, but not ESP and TFP, activity in crude extracts of bacteria expressing the respective proteins (Supplemental Fig. S4; data not shown). Thus, all of these genes encode nitrile specifier proteins, and we have named these proteins AtNSP2 (*At2g33070*), AtNSP3 (*At3g16390*), AtNSP4 (*At3g16410*), and AtNSP5 (*At5g48180*). Bacterial extracts harboring the *At3g07720* expression construct contained the recombinant protein, as demonstrated by western-blot analysis. However, neither the crude extract nor the purified Strep-tagged protein for *At3g07720* showed any specifier activity in our assays.

A Role for *At3g16400* in Simple Nitrile Formation Is Supported by QTL Analysis

We previously analyzed simple nitrile formation during glucosinolate hydrolysis in mature rosettes of 400 lines of the Bay-0 × Sha recombinant inbred line (RIL) population (Wentzell et al., 2008). This identified a QTL altering simple nitrile formation near *At3g16400*, but this is also in proximity to the previously cloned *ESM1/At3g14210* QTL. To test if this QTL may actually consist of multiple loci, we utilized the multiple interval mapping (MIM) algorithm of QTL Cartographer to further dissect the region. This showed that chromosome III has at least four QTLs for simple nitrile formation, one of which strongly associates with the physical location of *ESM1/At3g14210* (Fig. 8). MIM QTL analysis of expression data from the Bay-0 × Sha RIL population showed that there is a cis-eQTL for *At3g14210/ESM1*, likely from a promoter polymorphism (Zhang et al., 2006; West et al., 2007). The second QTL for simple nitrile formation on this chromosome lies directly on top of the physical location of the *At3g16390*, *At3g16400*, and *At3g16410* cluster. The Affymetrix ATH1 array probe set for *At3g16390*, *At3g16400*, and *At3g16410* also identifies a cis-eQTL at this region (Fig. 8). Thus, at least one of these genes is differentially expressed in the Bay-0 × Sha RIL population. Interestingly, the bottom of chromosome III had another QTL for simple nitrile formation that also was

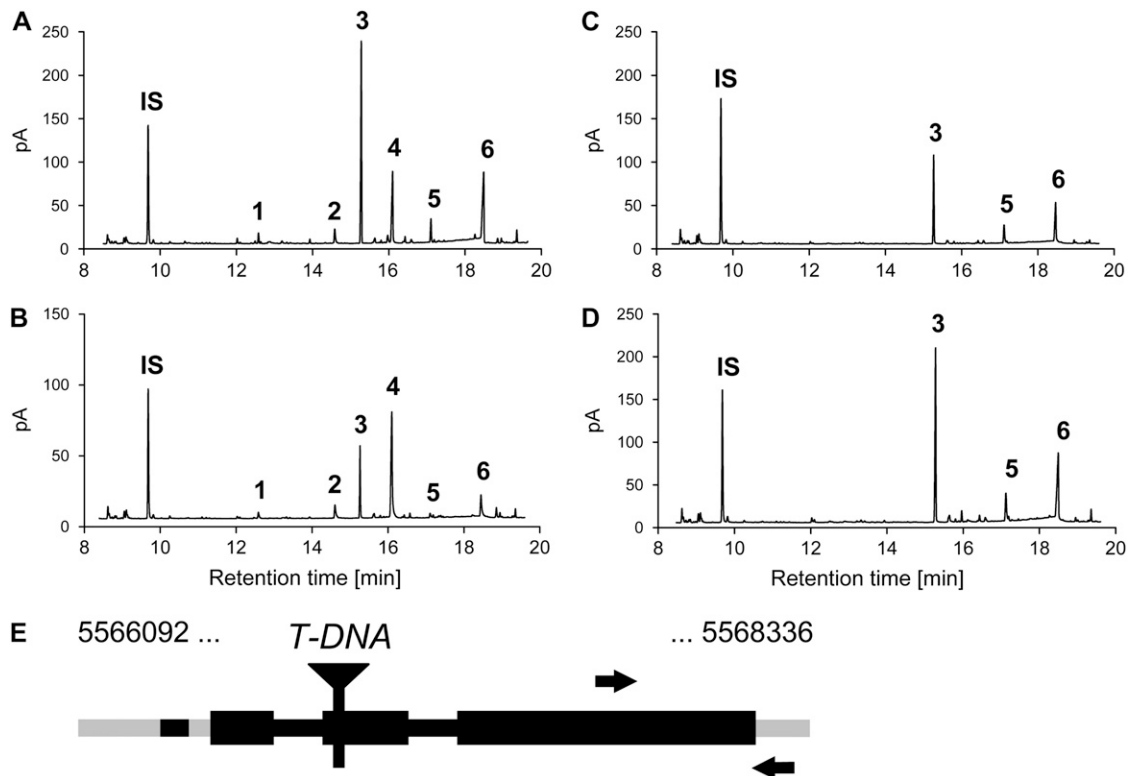


Figure 5. SALK_072600 is defective in constitutive and herbivore-induced simple nitrile formation. A to D, Glucosinolate hydrolysis products in homogenates of rosette leaves of *Arabidopsis* Col-0 wild type (A and B) or SALK_072600 (C and D). Leaves were harvested from control plants (no herbivore; A and C) or from plants infested with *P. rapae* larvae (B and D). Depicted are GC-FID traces. Peak 1, 5-Methylsulfanyl-pentanenitrile (nitrile product of 4-methylthiobutylglucosinolate); peak 2, 4-methanesulfinyl-butyronitrile (nitrile product of 3-methylsulfinylpropylglucosinolate); peak 3, 1-isothiocyanato-4-methylsulfanyl-butane (isothiocyanate product of 4-methylthiobutylglucosinolate); peak 4, 5-methanesulfinyl-pentanenitrile (nitrile product of 4-methylsulfinylbutylglucosinolate); peak 5, 1-isothiocyanato-3-methanesulfinyl-propane (isothiocyanate product of 3-methylsulfinylpropylglucosinolate); peak 6, 1-isothiocyanato-4-methanesulfinyl-butane (isothiocyanate product of 4-methylsulfinylbutylglucosinolate); peak IS, internal standard. E, Scheme of the *AtNSP1* gene (gene model *At3g16400.1*; www.Arabidopsis.org). Black boxes represent coding regions, and black and gray lines indicate introns and untranslated regions, respectively. The SALK_072600 line carries a T-DNA insertion in the second coding region. Arrows indicate the positions of the primers used for RT-PCR.

associated with altered expression of the *At3g16400* probe set (Fig. 8). As this region is in trans, it is likely a regulatory locus and provides an additional link between simple nitrile formation and the *At3g16390*, *At3g16400*, and *At3g16410* cluster. Thus, QTL and eQTL analysis provides further support for an in planta role of the *At3g16390*, *At3g16400*, and *At3g16410* cluster with simple nitrile formation in *Arabidopsis*.

Plant Specifier Protein Phylogenetic Analysis

To further refine the evolution of this family of specifier proteins, we obtained all protein sequences showing a significant BLASTP score across the length of the protein in comparison with *Arabidopsis* ESPs from the complete genomic sequences of several fungal, algal, bryophyte, and higher plant species that do not contain glucosinolates in addition to the known ESP

and TFP sequences. An *Arabidopsis* gene responsible for simple nitrile formation should not be closely related to genes from nonglucosinolate species. As expected from the biochemical assays, we found that *At3g16390*, *At3g16400*, *At3g16410*, *At2g33070*, and *At5g48180* are genes that do not have homologues in poplar (*Populus* species), rice (*Oryza sativa*), or grape (*Vitis vinifera*), species that do not contain glucosinolates (Fig. 9). In contrast, *At3g07720* appears to be a single-copy gene in *Arabidopsis* that has homologues in all fungal, algal, bryophyte, and higher plant species tested (Fig. 9). This suggests that *At3g07720* may encode the ancestral function from which the glucosinolate-related activity is derived.

Interestingly, AtNSP1 (*At3g16400*), AtNSP2 (*At2g33070*), AtNSP3 (*At3g16390*), and AtNSP4 (*At3g16410*) contain jacalin domains that are not present in AtNSP5 (*At5g48180*) or ESP or in the ancestral

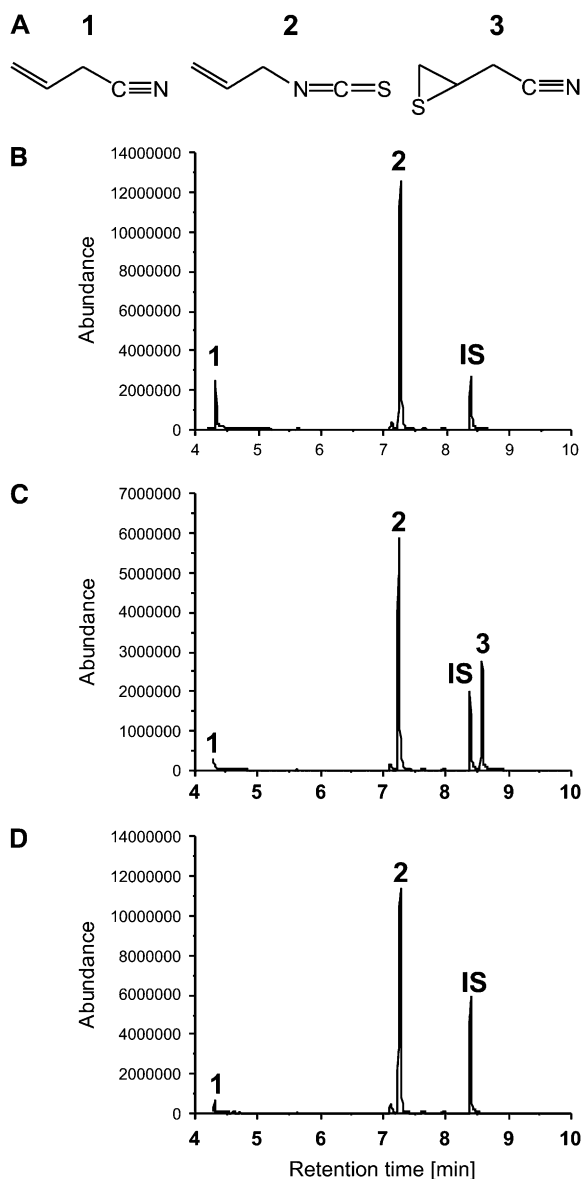


Figure 6. Effects of AtNSP1 on the hydrolysis of allylglucosinolate in vitro. A, Chemical structures of hydrolysis products formed from allylglucosinolate: 1, simple nitrile; 2, isothiocyanate; 3, epithionitrile. B to D, Hydrolysis products in enzyme assays carried out with purified AtNSP1 (B), purified ESP (C), or myrosinase alone (D) in 50 mM MES buffer, pH 6.0, containing 2 mM allylglucosinolate. Depicted are GC-MS chromatograms (total ion current traces) of dichloromethane extracts. IS, Internal standard; peak numbers refer to the structure numbers in A.

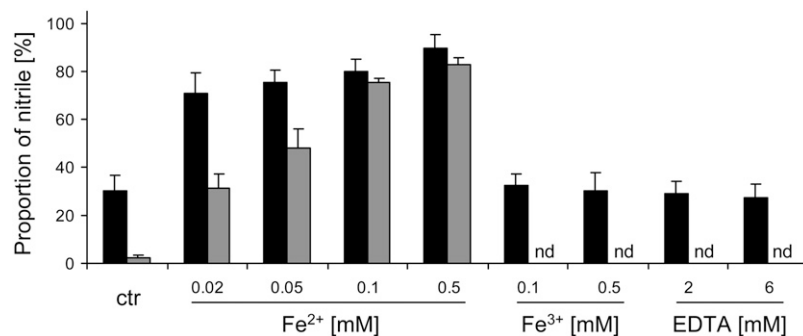
protein encoded by *At3g07720* (Fig. 4). This suggests that the sequences encoding these jacalin domains were obtained from some other gene. Phylogenetic analysis with the jacalin domains showed that they appear to be derived from the putative MBPs (Fig. 10). Interestingly, *At3g16410* seems to have obtained sequences for jacalin domains from two different sources (Fig. 10, compare *At3g16410a* and -b).

DISCUSSION

The chemical diversity of the glucosinolate-myrosinase system arises from variation in both glucosinolate biosynthesis and hydrolysis. As an example, a single protein amino acid, Met, is the precursor of more than 30 aliphatic glucosinolates within a single plant species, *Arabidopsis* (Reichelt et al., 2002). Hydrolysis of one of these glucosinolates, allylglucosinolate, can yield four product types with different biological activities, namely an isothiocyanate, a simple nitrile, an epithionitrile, and/or an organic thiocyanate, depending on the plant species and organ (Wittstock et al., 2003; Fig. 1). Although most studies have focused on the highly reactive and biologically active isothiocyanates, evidence is accumulating that simple nitriles are involved in a complex network of interactions between plants and their biotic environment (Wittstock et al., 2003; De Vos et al., 2008; Mumm et al., 2008). However, little is known about how simple nitrile formation is regulated at the genetic or biochemical level. In this report, we present evidence that simple nitrile formation upon myrosinase-catalyzed glucosinolate hydrolysis in *Arabidopsis* is controlled by a set of five ESP homologues, NSP1 through NSP5, that do not support epithionitrile formation. These homologues are encoded by a small gene family that also includes the *ESP* gene responsible for epithionitrile formation as a sixth member. Our approach was based on the hypothesis that a Kelch repeat protein with similarity to ESP is involved in simple nitrile formation. Among six candidate genes in *Arabidopsis*, we identify *AtNSP1* (*At3g16400*) as the gene responsible for constitutive and herbivore-induced simple nitrile formation in Col-0 rosette leaves, based on the phenotype of a knockout mutant of this gene (SALK_072600), the regulation of its transcription in rosette leaves, and the ability of the encoded protein to promote simple nitrile, but not epithionitrile, formation in vitro. Identification of *AtNSP1* is further supported by QTL analysis, which collocated a gene cluster including this gene with a QTL for simple nitrile formation. Heterologous protein expression shows that another four of the six phylogenetically identified candidate genes encode NSPs, while the sixth gene likely represents the ancestor of the gene family with a different function.

Our study demonstrates that *Arabidopsis* is equipped with an extensive protein machinery dedicated to simple nitrile formation. This is in agreement with the observation that simple nitriles are major glucosinolate hydrolysis products in certain organs and developmental stages of *Arabidopsis* (Wentzell and Kliebenstein, 2008) and that their formation is induced upon herbivory (Fig. 2). The phenotype of the knockout mutant allows us to exclude a major contribution of the *P. rapae* NSP to increased simple nitrile formation within *Arabidopsis* rosettes upon herbivory. Thus, simple nitriles might fulfill important functions for the plant. Recent studies indicate that these functions are likely linked to the volatility of the simple nitriles.

Figure 7. Effects of iron salts and EDTA on AtNSP1 in vitro. Simple nitrile formation in assays with purified AtNSP1 and myrosinase (black bars) and in control assays with myrosinase alone (gray bars) was measured in 50 mM MES, pH 6.0, containing 2 mM allylglucosinolate. Fe²⁺ and Fe³⁺ were added as (NH₄)₂[Fe(SO₄)₂] and FeCl₃, respectively. Means ± SD of results obtained in three independent experiments are shown. ctr, Control without addition of iron; nd, not determined.



P. rapae-infested Col-0 plants engineered to produce a higher proportion of simple nitriles due to overexpression of the *ESP* cDNA were significantly more attractive to the parasitoid wasp *Cotesia rubecula* than infested wild-type plants (Mumm et al., 2008). Additionally, the increased proportion of nitriles formed in the *ESP*-overexpressing plants or the application of indol-3-acetonitrile to an *Arabidopsis* mutant deficient in indole glucosinolate biosynthesis decreased the number of eggs laid by *P. rapae* females (De Vos et al., 2008; Mumm et al., 2008). Simple nitrile formation has been shown to be associated with higher damage by generalist herbivores, likely due to impaired direct defense by isothiocyanates (Lambrix et al., 2001; Burow et al., 2006b; Zhang et al., 2006). This effect was detectable even upon a minor increase in simple nitrile formation (Zhang et al., 2006). In contrast, the isothiocyanate, but not the simple nitrile from 3-butenylglucosinolate, elicited the strongest electroantennogram responses in the aphid parasitoid *Diaeretiella rapae* (Pope et al., 2008). Thus, the ratio of simple nitrile to isothiocyanate formed in glucosinolate-containing plants appears to have complex, opposing, and multitrophic effects on the insect community. This suggests that a plant may benefit from tight regulation of simple nitrile formation to maximize its direct versus indirect defense (Lambrix et al., 2001; Burow et al., 2006b; Zhang et al., 2006). In our experiments, induction of simple nitrile formation was a local effect detected only in the *P. rapae*-damaged leaves (Fig. 2). Such a local response may be beneficial, first, because it may help parasitoids to locate their host on the plant, and second, because it restricts the decrease in directly defensive isothiocyanates to a limited area.

AtNSP1 through -5 are the first examples of plant proteins that promote simple nitrile formation but do not support epithionitrile formation from alkenylglucosinolates (Fig. 6; Supplemental Figs. S2 and S4). In contrast, *ESP* allows simple nitrile formation only from glucosinolates that cannot form epithionitriles due to their lack of a terminal double bond. During myrosinase-catalyzed hydrolysis in the presence of *ESP*, glucosinolates with a terminal double bond in their side chain undergo an intramolecular sulfur migration in which the thioglucosidic sulfur is transferred to the double bond, yielding a thiirane ring (Brocker and Benn, 1983). Previous in vitro studies have shown that while the formation of epithionitriles is strictly

dependent on *ESP*, simple nitrile formation can also be enhanced nonenzymatically at pH < 3 (Ettlinger et al., 1961) and in the presence of Fe²⁺ (for review, see Wittstock and Burow, 2007). Taking this into account, the presence of five *NSP* genes in *Arabidopsis* indicates that (1) the effects of low pH and Fe²⁺ in the absence of

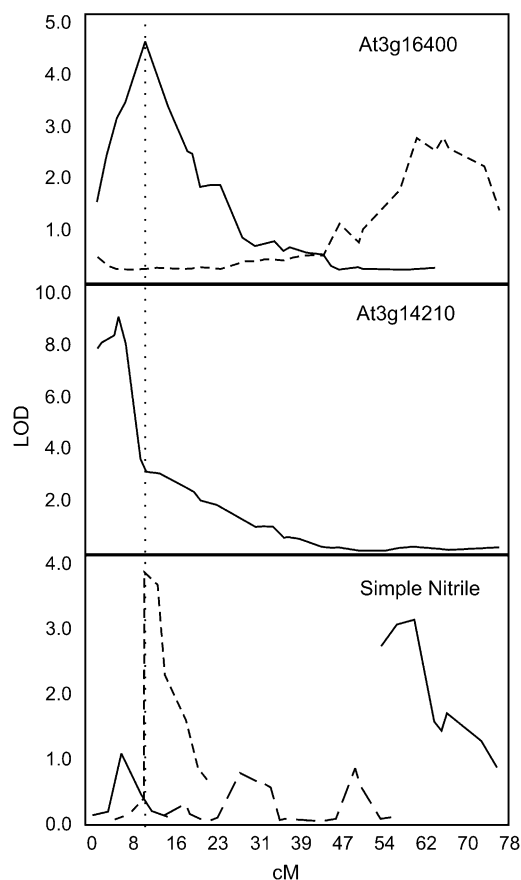


Figure 8. Expression and phenotypic QTLs in the *Arabidopsis* Bay-0 × Sha RIL population. QTLs on chromosome III governing the expression of *AtNSP1* (*At3g16400*) and *ESM1* (*At3g14210*) and the formation of simple nitriles are shown. Different line styles (dashed, solid, etc.) show independent QTLs for each phenotype and the QTL significant conditional likelihood profiles as identified by the MIM module of QTL Cartographer. The vertical line through the three graphs shows the physical position of *At3g16400* as determined by high-resolution mapping. This analysis focused only on chromosome III for clarity and precision. cM, Centimorgan position along chromosome III; LOD, logarithm of the odds.

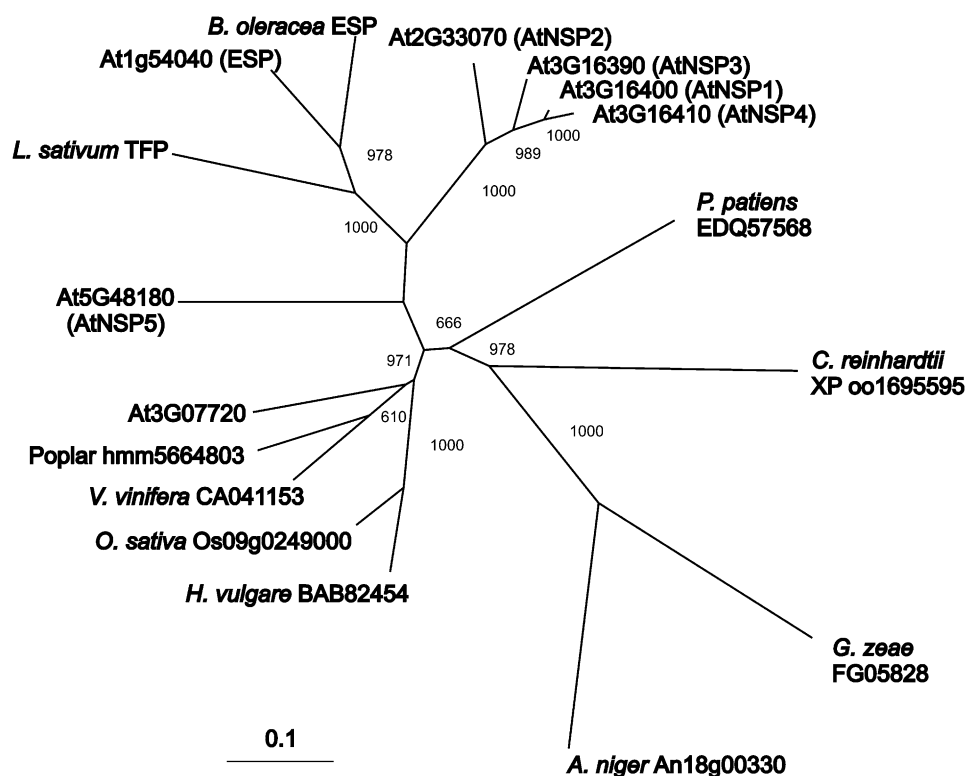


Figure 9. Phylogeny of the Kelch domain. For all higher plant ESP homologues as well as a bryophyte, an algal, and two fungal homologues, the portion of each protein sequence containing the Kelch domains was used to generate a phylogeny. The neighbor-joining tree is shown with bootstrap values over 600 out of 1,000. A distance scale is included at the bottom for the protein tree. An unrooted tree is presented to allow a focus on relatedness. Source organisms were *Arabidopsis*, *Populus trichocarpa*, *Vitis vinifera*, *Oryza sativa*, *Hordeum vulgare*, *Aspergillus niger*, *Gibberella zeae*, *Chlamydomonas reinhardtii*, *Physcomitrella patiens*, *Lepidium sativum*, and *Brassica oleracea*.

specifier proteins are not relevant as a means of generating simple nitriles under physiological conditions, (2) the potential need to tightly regulate nitrile formation may require tools that can be differentially expressed in time and space, and (3) there may be a biological need to produce simple nitriles from glucosinolates with a terminal double bond in their side chains. The broad substrate specificity of AtNSP1 suggests that there is a general function of simple nitriles. The biochemical role of specifier proteins in simple nitrile formation is unknown, but it is likely the abstraction of the thioglucosidic sulfur from the aglucone (Wittstock and Burow, 2007). A similar role has been proposed for Fe^{2+} in simple nitrile formation in the absence of specifier proteins (Foo et al., 2000). In our experiments with AtNSP1, Fe^{2+} did not affect NSP-dependent simple nitrile formation apart from its strong effect on the background simple nitrile levels, and NSP activity did not require the presence of iron (Fig. 7; Supplemental Fig. S3). This is in agreement with previous results obtained using PrNSP, the NSP isolated from larvae of *P. rapae*. Thus, in contrast to epithionitrile formation by ESP, which is strictly dependent on iron, simple nitrile formation by specifier proteins is not iron dependent.

As the known ESPs, AtNSP1 through -5 are Kelch repeat proteins containing four or five Kelch domains (Fig. 4). In contrast to the ESPs and AtNSP5, four of the NSPs possess one (AtNSP1, AtNSP2, and AtNSP3) or two (AtNSP4) lectin-like jacalin domains at their N terminus with high amino acid sequence similarity to the putative Arabidopsis MBPs. Despite these remarkable structural differences, all five NSPs have simple

nitrile-forming but no epithionitrile- or thiocyanate-forming activity. MBPs have long been associated with the glucosinolate-myrosinase system (Falk et al., 1995; Taipalensuu et al., 1997; Zhang et al., 2006). In *Brassica* species, they have been shown to form stable complexes with myrosinases (Falk et al., 1995; Geshi and Brandt, 1998). Their biological functions, however, are unknown. Antisense suppression of MBP in transgenic *Brassica napus* prevented myrosinase complex formation but did not affect the hydrolysis product profile of exogenous 4-hydroxybenzylglucosinolate (Eriksson et al., 2002). Furthermore, the performance and consumption rates, respectively, of a generalist and a specialist insect herbivore as well as the growth rates of two fungal pathogens were unaffected (Eriksson et al., 2002). Based on the similarity of the jacalin domains to MBPs and the absence of Kelch domains in MBPs, a putative function of the jacalin domains might be to affect the interaction of the NSPs with myrosinase. Given that ESP and AtNSP5 are perfectly capable of acting as specifier proteins during glucosinolate hydrolysis even though they lack a jacalin domain, this domain may rather affect other NSP characteristics, such as tissue localization and regulation. Taken together, it remains to be shown if and how the Kelch domains and jacalin-like domains contribute to the catalytic activity of the NSPs or to the functioning of the proteins in planta.

As At3g07720 is the basal member of the family and lacks a jacalin domain, the presence of jacalin domains in AtNSP1 through -4 is likely the derived state. Interestingly, phylogenetic analysis suggests that these do-

the organisms interacting with them at different trophic levels.

MATERIALS AND METHODS

Plants and Insects

Arabidopsis (*Arabidopsis thaliana*) plants of the accession Col-0 were grown on soil in a controlled-environment chamber at 22°C, 60% to 70% relative humidity, and 300 $\mu\text{mol m}^{-2} \text{s}^{-1}$ photosynthetically active radiation. The photoperiod was 10 h. Seeds of the T-DNA insertion line SALK_072600 and all other mutant lines (Col-0 background; Supplemental Table S3) were obtained from the Nottingham Arabidopsis Stock Centre and grown on soil under the above conditions or germinated on Murashige and Skoog medium containing 0.9% (w/v) agar and 50 $\mu\text{g mL}^{-1}$ kanamycin and transferred to soil after 14 d. A continuous rearing of *Pieris rapae* (Lepidoptera, Pieridae) was maintained on brussels sprout (*Brassica oleracea gemmifera*) plants in a controlled-environment chamber at 24°C, 60% relative humidity, and a photoperiod of 16 h.

Chemicals

Allylglucosinolate was purchased from AppliChem. All other intact glucosinolates were isolated as described previously (Thies, 1988; Zrybko et al., 1997). Allyl isothiocyanate (3-isothiocyanatoprop-1-ene) and allyl cyanide (but-3-enitrile) standards were obtained from Fluka; benzyl isothiocyanate [1-(isothiocyanatomethyl) benzene] and benzyl cyanide (phenylacetone nitrile) were procured from Sigma-Aldrich; benzonitrile was from Merck.

Gas Chromatography Analysis of Glucosinolate Hydrolysis Products

Crude plant extracts and ESP/NSP assays were extracted with dichloromethane and analyzed by gas chromatography-mass spectrometry (GC-MS) and by gas chromatography-flame ionization detection (GC-FID) using an Agilent 6890N series gas chromatograph with an HP5MS column (30 m \times 0.25 mm \times 0.25 μm ; Wicom), splitless injection at 200°C (injection volume, 1 μL), and the temperature programs listed in Supplemental Table S3. MS, FID, and quantification by GC-FID were carried out as described previously (Lambrix et al., 2001). Products were identified by comparison of mass spectra and retention times with those of authentic standards and with published MS spectra (Spencer and Daxenbichler, 1980).

Plant Extracts for Analyses of Hydrolysis Products from Endogenous Glucosinolates and Myrosinase Activity Assays

Crude plant extracts were prepared by grinding 150 to 500 mg (fresh weight) of rosette leaves with 50 mM MES buffer, pH 6.0, or 50 mM Tris-HCl, pH 7.5 (1 mL of buffer per gram). After incubation at room temperature for 5 min, samples were centrifuged at 10,000g and 4°C for 10 min. The supernatants were kept on ice until used as described below. Protein determination was done according to Bradford (1976) using bovine serum albumin as a standard.

Analysis of Glucosinolate Hydrolysis Products from Endogenous Glucosinolates in Plant Extracts

Crude plant extracts (150–350 μL) to which 30 to 50 μL of benzonitrile (100 $\mu\text{g mL}^{-1}$ in methanol) as an internal standard had been added were extracted with 2 \times 750 μL of dichloromethane, and the organic phases were combined, dried over Na_2SO_4 , concentrated under an air stream, and analyzed by GC-FID and GC-MS as described above.

Myrosinase Assay

Myrosinase assays were conducted with 2 mM allylglucosinolate and 30 μL of plant extract in a total volume of 100 μL . After incubation at 37°C for 15 min, the reaction was stopped by boiling (95°C for 5 min). Glc concentrations were determined as described previously (Burow et al., 2006b).

Analysis of Glucosinolate Hydrolysis Products from Allylglucosinolate in Plant Extracts

Rosette leaves (170–260 mg fresh weight) were ground in 300 μL of 50 mM MES buffer, pH 6.0. After addition of 700 μL of MES buffer and incubation with 10 to 40 μL of 10 mM allylglucosinolate at room temperature for 5 min, 50 μL of benzonitrile (100 $\mu\text{g mL}^{-1}$ in methanol) was added as an internal standard (Wentzell and Kliebenstein, 2008). Glucosinolate hydrolysis products were extracted with 2 \times 2 mL of dichloromethane and prepared for GC-FID analysis as described above.

Glucosinolate Analysis

Preparation of samples for glucosinolate determination as desulfoglucosinolates was as described previously (Burow et al., 2006b). Twenty to 50 μL of the samples was analyzed by HPLC on an Agilent HP1200 Series instrument equipped with a C-18 reverse-phase column (LiChrospher 100 RP18, 250 \times 4.6 mm, 5- μm particle size; Wicom) as described (Burow et al., 2006b). Desulfoglucosinolates were identified based on comparison of retention times and UV absorption spectra with those of known standards (Reichelt et al., 2002), and glucosinolate contents were calculated as described (Burow et al., 2006b).

Herbivore Induction

P. rapae larvae (second and third instar) were placed on rosettes of 6-week-old Arabidopsis Col-0 wild-type or SALK_072600 plants (one larva per plant). Plants were covered with plastic bread bags to prevent insects from escaping and placed in a controlled-environment chamber (14 h of light/10 h of dark, 22°C, 60% relative humidity). Untreated control plants were kept in bags under the same conditions. After 24 h, leaves were harvested for expression analysis by RT-PCR and for analysis of glucosinolate hydrolysis products. From herbivore-treated plants, damaged and undamaged leaves were harvested separately.

Identification of Candidate Genes

To identify candidate genes encoding a nitrile-specifier protein in Arabidopsis Col-0, we used the Arabidopsis (*Lea*) ESP full-length cDNA (1,026 bp; GenBank accession no. AF416787) as a query to search the Arabidopsis genome using the BLASTN algorithm (<http://www.ncbi.nlm.nih.gov/blast/Blast.cgi>). Full-length sequences with 35% or greater identity with the ESP cDNA were translated in silico, and protein structures were predicted using InterProScan (<http://www.ebi.ac.uk/InterProScan/>; European Bioinformatics Institute). Genes encoding Kelch repeat proteins were chosen for further analyses.

Verification of the T-DNA Insertion

Genomic DNA was isolated from frozen leaf material using the plantDNA-OLS-Kit (OMNI Life Science) according to the manufacturer's instructions. PCR was conducted with Taq polymerase (OMNI Life Science) in 25 μL of PCR buffer containing 1 unit of enzyme, 200 ng of genomic DNA, 200 μM deoxy-nucleoside triphosphates (dNTPs), and 20 pmol of each primer (gene-specific primer for *At3g16400*, 5'-CCTACTTGTGCTATGGGATGTGAG-3'; T-DNA left border primer, 5'-CGTGGACCGCTTGCTGC-3') as detailed in Supplemental Table S4. PCR results were analyzed on 1% agarose gels, and the amplified fragment was verified by sequencing.

Expression Analysis by RT-PCR

Total RNA was isolated from frozen plant tissues using ribozol-OLS (OMNI Life Science) as directed and quantified spectrophotometrically. To ensure that data points within the linear range of the RT-PCR were obtained, first-strand cDNA synthesis was performed on three different amounts of RNA (0.1, 0.3, and 1.0 μg). RT reactions were carried out with the totalscript-OLS kit (OMNI Life Science) as directed and purified using the Wizard SV Gel and PCR Clean-Up System (Promega). PCR was done in a total volume of 25 μL of PCR buffer containing 1 unit of omnitaq-OLS (OMNI Life Science), 0.5 μL of the RT reaction, 200 μM dNTPs, and 10 pmol of each primer (Supplemental Table S5) as detailed in Supplemental Table S4. Twenty microliters of the PCR product was analyzed on 1.5% agarose gels. The 604-bp fragment amplified with the *AtNSP1* primers was cloned and verified by sequencing.

Generation of Expression Constructs

Total RNA was isolated from rosette leaves using the RNeasy Plant Mini kit (Qiagen) and from 7-d-old seedlings using the plant RNA-OLS kit (OMNI Life Science) as directed. cDNA synthesis was done using SuperScript III reverse transcriptase (Invitrogen) and oligo(dT)₁₂₋₁₈ primers (Invitrogen) according to the manufacturer's directions. The full-length cDNAs of *At3g07720* (990 bp) and *At5g48180* (981 bp) were cloned from seedling cDNA, and those of *At2g33070* (1,416 bp) and *At3g16390* (1,404 bp) were cloned from leaf cDNA. The full-length cDNAs of *At3g16400* (1,413 bp) and *At3g16410* (1,860 bp) were amplified from the ABRC cDNA clones U09267 and S81471, respectively (ABRC DNA Stock Center). Primers are listed in Supplemental Table S6. The Arabidopsis *Ler* ESP cDNA was amplified from a previously described ESP construct in the pCRT7/CT-TOPO vector (Invitrogen; Burow et al., 2006a). PCR was done in a total volume of 25 μ L of PCR buffer containing 1.25 units of PfuTurbo C_x Hotstart DNA polymerase (Stratagene), 20 to 200 ng of cDNA or 20 ng of plasmid DNA, 200 μ M dNTPs, and 10 to 20 pmol of each primer as detailed in Supplemental Table S4. PCR products were cloned into pET52(b)+, which had been modified and prepared for USER cloning as described (Nour-Eldin et al., 2006) and verified by sequencing.

Heterologous Expression and Protein Purification

Escherichia coli strain BL21(DE3)pLysS (Invitrogen) was transformed with the empty expression vector [modified pET52(b)+] or one of the expression constructs. Single colonies selected on Luria-Bertani medium supplemented with 100 μ g mL⁻¹ ampicillin and 34 μ g mL⁻¹ chloramphenicol were used to inoculate 10 mL of terrific broth medium containing the same antibiotics. After growth at 18°C and 220 rpm for 62 h, 6 mL of the precultures was transferred to 1 L of terrific broth medium with antibiotics. The cultures were grown and extracted as described (Burow et al., 2006a). The crude bacterial extract was directly used in ESP/NSP activity assays or loaded onto Strep-Tactin Sepharose resin (IBA) to purify the recombinant proteins by affinity chromatography. Deviating from the manufacturer's directions, no EDTA was added to the buffer used for elution of Strep-tagged proteins. The purity of the fractions was analyzed by Tris/SDS-PAGE. Protein content was determined as described above.

Western-Blot Analysis

Crude bacterial extracts or fractions from protein purifications were separated on 12% Tris/SDS-PAGE gels and transferred to nitrocellulose membranes. Strep-tagged proteins were detected using an anti-Strep tag antibody (Strep-Tactin alkaline phosphatase conjugate; IBA) according to the manufacturer's instructions.

ESP/NSP Activity Assays with Recombinant Proteins

Enzyme assays were carried out in 500 μ L of 50 mM MES buffer, pH 6.0, with 150 μ L of bacterial extracts or 1 unit of purified recombinant protein (unless otherwise stated). Glucosinolates were added to the assay mixtures to a final concentration of 1 mM (4-methylthiobutyl- and benzylglucosinolate), 2 mM (allylglucosinolate), or 3 mM (4-methylsulfinylbutylglucosinolate). The reaction was started by the addition of 4 units of myrosinase purified from white mustard (*Sinapis alba*) seeds (Burow et al., 2006a). After incubation at room temperature for 40 min, 50 μ L of benzonitrile (100 ng μ L⁻¹ in methanol) was added as an internal standard. The assay mixtures were extracted with dichloromethane (see above) and analyzed by GC-MS and GC-FID. In fractions of independent purifications, 1 unit corresponded to 20 to 25 μ g of purified AtNSP1 or to 5 to 10 μ g of ESP.

Analysis of QTLs for Gene Expression and Glucosinolate Hydrolysis

Transcript accumulation values for *At3g14210* and *At3g16390/400/410* were obtained from previous analysis of the Bay-0 \times Sha population (West et al., 2007). Average values of the proportion of simple nitriles produced during glucosinolate hydrolysis were obtained from a separate analysis of the same population (Wentzell et al., 2008). To precisely map QTLs, we utilized MIM within Windows QTL Cartographer version 2.0 (Zeng et al., 1999). MIM employs forward and backward regression to generate an initial model. This

model was subjected to at least three rounds of QTL position optimization, a new QTL search, and an epistasis search among detected QTLs before QTLs were tested for inclusion in the final model within the MIM algorithm. This was conducted separately for the two expression traits and one hydrolysis trait. Only the results from chromosome III are presented for clarity and focus.

Phylogenetic Analysis

We used BLASTP to query plant, fungal, bryophyte, and algal whole genome sequences with the full-length ESP amino acid sequence. To include a gene as a potential ESP homologue, it had to have a BLASTP score of at least e^{-45} with similarity across the full ESP sequence. The contiguous Kelch domains of the proteins were subjected to ClustalX phylogenetic analysis using neighbor joining with 1,000 bootstraps. To compare the evolutionary history of the jacalin domains present in the Arabidopsis specifier proteins, we identified all homologous proteins encoded in the rice (*Oryza sativa*) and poplar (*Populus* species) genomes as described above using the jacalin domain from Arabidopsis MBP1 (*At1g52040*). We limited our analysis to these three genomes because of the large number of jacalin homologues. We then conducted ClustalX phylogenetic analysis using neighbor joining with 1,000 bootstraps. The tree was visualized in Treeview as an unrooted tree, given the fact that we do not know the direction of evolution or if it is monodirectional for these families. Instead, we focused on relatedness of the sequences. Maximum likelihood trees were similar in structure.

Supplemental Data

The following materials are available in the online version of this article.

Supplemental Figure S1. Myrosinase activity in Col-0 and SALK_072600 rosette leaves.

Supplemental Figure S2. Effects of AtNSP1 on glucosinolate hydrolysis in vitro.

Supplemental Figure S3. Effects of iron salts and EDTA on AtNSP1 activity with various substrates in vitro.

Supplemental Figure S4. Effects of AtNSP2 through -5 on allylglucosinolate hydrolysis in vitro.

Supplemental Table S1. Publicly available mutant lines of candidate genes screened in this study.

Supplemental Table S2. Glucosinolate contents in Col-0 and SALK_072600 rosette leaves.

Supplemental Table S3. Temperature programs for GC-MS and GC-FID analyses.

Supplemental Table S4. Temperature programs used for PCR.

Supplemental Table S5. Primers for expression analysis by RT-PCR.

Supplemental Table S6. Primers for the generation of expression constructs.

ACKNOWLEDGMENTS

We thank Claudine Theuring for isolating glucosinolate substrates, Anita Backenköhler and Loretta Heise for excellent technical assistance, and the ABRC and Nottingham Arabidopsis Stock Centre for providing EST clones and seed stocks, respectively.

Received October 2, 2008; accepted October 31, 2008; published November 5, 2008.

LITERATURE CITED

- Adams J, Kelso R, Cooley L (2000) The Kelch repeat superfamily of proteins: propellers of cell function. *Trends Cell Biol* **10**: 17–24
- Agrawal AA, Kurashige NS (2003) A role for isothiocyanates in plant resistance against the specialist herbivore *Pieris rapae*. *J Chem Ecol* **29**: 1403–1415

- Andréasson E, Jørgensen LB (2003) Localization of plant myrosinases and glucosinolates. In JT Romeo, ed, *Integrative Phytochemistry: From Ethnobotany to Molecular Ecology*, Vol 37. Elsevier, Amsterdam, pp 79–99
- Benderoth M, Textor S, Windsor AJ, Mitchell-Olds T, Gershenzon J, Kroymann J (2006) Positive selection driving diversification in plant secondary metabolism. *Proc Natl Acad Sci USA* **103**: 9118–9123
- Bennett RN, Rosa EA, Mellon FA, Kroon PA (2006) Ontogenic profiling of glucosinolates, flavonoids, and other secondary metabolites in *Eruca sativa* (salad rocket), *Diplotaxis erucoides* (wall rocket), *Diplotaxis tenuifolia* (wild rocket), and *Bunias orientalis* (Turkish rocket). *J Agric Food Chem* **54**: 4005–4015
- Bradford MM (1976) A rapid and sensitive method for the quantification of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**: 248–254
- Brocker ER, Benn MH (1983) The intramolecular formation of epithioalkenenitriles from alkenylglucosinolates by *Crambe abyssinica* seed flour. *Phytochemistry* **22**: 770–772
- Burow M, Bergner A, Gershenzon J, Wittstock U (2007a) Glucosinolate hydrolysis in *Lepidium sativum*: identification of the thiocyanate-forming protein. *Plant Mol Biol* **63**: 49–61
- Burow M, Markert J, Gershenzon J, Wittstock U (2006a) Comparative biochemical characterization of nitrile-forming proteins from plants and insects that alter myrosinase-catalysed hydrolysis of glucosinolates. *FEBS J* **273**: 2432–2446
- Burow M, Müller R, Gershenzon J, Wittstock U (2006b) Altered glucosinolate hydrolysis in genetically engineered *Arabidopsis thaliana* and its influence on the larval development of *Spodoptera littoralis*. *J Chem Ecol* **32**: 2333–2349
- Burow M, Rice M, Hause B, Gershenzon J, Wittstock U (2007b) Cell- and tissue-specific localization and regulation of the epithiospecifier protein in *Arabidopsis thaliana*. *Plant Mol Biol* **64**: 173–185
- Burow M, Zhang ZY, Ober JA, Lambrix VM, Wittstock U, Gershenzon J, Kliebenstein DJ (2008) ESP and ESM1 mediate indol-3-acetonitrile production from indol-3-ylmethyl glucosinolate in Arabidopsis. *Phytochemistry* **69**: 663–671
- Chew FS (1988) Biological effects of glucosinolates. In HG Cutler, ed, *Biologically Active Natural Products*. American Chemical Society, Washington, DC, pp 155–181
- De Vos M, Kriksunov KL, Jander G (2008) Indole-3-acetonitrile production from indole glucosinolates deters oviposition by *Pieris rapae* (white cabbage butterfly). *Plant Physiol* **146**: 916–926
- Eriksson S, Andréasson E, Ekblom B, Graner G, Pontoppidan B, Taipalensuu J, Zhang JM, Rask L, Meijer J (2002) Complex formation of myrosinase isoenzymes in oilseed rape seeds are dependent on the presence of myrosinase-binding proteins. *Plant Physiol* **129**: 1592–1599
- Ettlinger MG, Dateo GP, Harrison BW, Mabry TJ, Thompson C (1961) Vitamin C as a coenzyme: the hydrolysis of mustard oil glucosides. *Proc Natl Acad Sci USA* **47**: 1875–1880
- Fahey JW, Zalcmann AT, Talalay P (2001) The chemical diversity and distribution of glucosinolates and isothiocyanates among plants. *Phytochemistry* **56**: 5–51
- Fahey JW, Zhang Y, Talalay P (1997) Broccoli sprouts: an exceptionally rich source of inducers of enzymes that protect against chemical carcinogens. *Proc Natl Acad Sci USA* **94**: 10367–10372
- Falk A, Taipalensuu J, Ek B, Lenman M, Rask L (1995) Characterization of rapeseed myrosinase-binding protein. *Planta* **195**: 387–395
- Foo HL, Groenning LM, Goodenough L, Bones AM, Danielsen BE, Whiting DA, Rossiter JT (2000) Purification and characterisation of epithiospecifier protein from *Brassica napus*: enzymic intramolecular sulphur addition within alkenyl thiohydroximates derived from alkenyl glucosinolate hydrolysis. *FEBS Lett* **468**: 243–246
- Geshi N, Brandt A (1998) Two jasmonate-inducible myrosinase-binding proteins from *Brassica napus* L. seedlings with homology to jacalin. *Planta* **204**: 295–304
- Halkier BA, Gershenzon J (2006) Biology and biochemistry of glucosinolates. *Annu Rev Plant Biol* **57**: 303–333
- Kessler A, Baldwin IT (2002) Plant responses to insect herbivory: the emerging molecular analysis. *Annu Rev Plant Physiol Plant Mol Biol* **53**: 299–328
- Kliebenstein DJ (2004) Secondary metabolites and plant/environment interactions: a view through *Arabidopsis thaliana* tinged glasses. *Plant Cell Environ* **27**: 675–684
- Kliebenstein DJ, Kroymann J, Mitchell-Olds T (2005) The glucosinolate-myrosinase system in an ecological and evolutionary context. *Curr Opin Plant Biol* **8**: 264–271
- Kost C, Heil M (2006) Herbivore-induced plant volatiles induce an indirect defence in neighbouring plants. *J Ecol* **94**: 619–628
- Lambrix V, Reichelt M, Mitchell-Olds T, Kliebenstein DJ, Gershenzon J (2001) The *Arabidopsis* epithiospecifier protein promotes the hydrolysis of glucosinolates to nitriles and influences *Trichoplusia ni* herbivory. *Plant Cell* **13**: 2793–2807
- Lazzeri L, Curto G, Leoni O, Dallavalle E (2004) Effects of glucosinolates and their enzymatic hydrolysis products via myrosinase on the root-knot nematode *Meloidogyne incognita* (Kofoid et White) Chitw. *J Agric Food Chem* **52**: 6703–6707
- Lichtenstein EP, Strong FM, Morgan DG (1962) Identification of 2-phenylethylthiocyanate as an insecticide occurring naturally in the edible part of turnips. *J Agric Food Chem* **10**: 30–33
- Matusheski NV, Swarup R, Juvik JA, Mithen R, Bennett M, Jeffery EH (2006) Epithiospecifier protein from broccoli (*Brassica oleracea* L. ssp. *italica*) inhibits formation of the anticancer agent sulforaphane. *J Agric Food Chem* **54**: 2069–2076
- Miao Y, Zentgraf U (2007) The antagonist function of *Arabidopsis* WRKY53 and ESR/ESP in leaf senescence is modulated by the jasmonic and salicylic acid equilibrium. *Plant Cell* **19**: 819–830
- Mumm R, Burow M, Bukovinszkyne K, Kazantzidou E, Wittstock U, Dicke M, Gershenzon J (2008) Formation of simple nitriles upon glucosinolate hydrolysis affects direct and indirect defense against the specialist herbivore, *Pieris rapae*. *J Chem Ecol* **34**: 1311–1321
- Nour-Eldin HH, Hansen BG, Norholm MH, Jensen JK, Halkier BA (2006) Advancing uracil-excision based cloning towards an ideal technique for cloning PCR fragments. *Nucleic Acids Res* **18**: e122
- Pieterse CMJ, Dicke M (2007) Plant interactions with microbes and insects: from molecular mechanisms to ecology. *Trends Plant Sci* **12**: 564–569
- Pope TW, Kissen R, Grant M, Pickett JA, Rossiter JT, Powell G (2008) Comparative innate responses of the aphid parasitoid *Diaeretiella rapae* to alkenylglucosinolate-derived isothiocyanates, nitriles, and epithionitriles. *J Chem Ecol* **34**: 1302–1310
- Rask L, Andréasson E, Ekblom B, Eriksson S, Pontoppidan B, Meijer J (2000) Myrosinase: gene family evolution and herbivore defense in Brassicaceae. *Plant Mol Biol* **42**: 93–113
- Reichelt M, Brown PD, Schneider B, Oldham NJ, Stauber E, Tokuhisa J, Kliebenstein DJ, Mitchell-Olds T, Gershenzon J (2002) Benzoic acid glucosinolate esters and other glucosinolates from *Arabidopsis thaliana*. *Phytochemistry* **59**: 663–671
- Spencer FG, Daxenbichler ME (1980) Gas chromatography-mass spectrometry of nitriles, isothiocyanates and oxazolidinethiones derived from cruciferous glucosinolates. *J Sci Food Agric* **31**: 359–367
- Taipalensuu J, Eriksson S, Rask L (1997) The myrosinase-binding protein from *Brassica napus* seeds possesses lectin activity and has a highly similar vegetatively expressed wound-inducible counterpart. *Eur J Biochem* **250**: 680–688
- Thies W (1988) Isolation of sinigrin and glucotropaeolin from cruciferous seeds. *FETT Wissenschaft Technologie-Fat Sci Technol* **90**: 311–314
- Tierens K, Thomma BPH, Brouwer M, Schmidt J, Kistner K, Porzel A, Mauch-Mani B, Cammue BPA, Broekaert WF (2001) Study of the role of antimicrobial glucosinolate-derived isothiocyanates in resistance of Arabidopsis to microbial pathogens. *Plant Physiol* **125**: 1688–1699
- Tookey HL (1973) Crambe thioglucoside glucohydrolase (EC 3.2.3.1): separation of a protein required for epithiobutane formation. *Can J Biochem* **51**: 1654–1660
- Van Poppel G, Verhoeven DTH, Verhagen H, Goldbohm RA (2000) Brassica vegetables and cancer prevention: epidemiology and mechanisms. *Adv Exp Med Biol* **472**: 159–168
- Wentzell AM, Boeye I, Zhang ZY, Kliebenstein D (2008) Genetic networks controlling structural outcome of glucosinolate activation across development. *PLoS Genetics* **4**: e1000234
- Wentzell AM, Kliebenstein D (2008) Genotype, age, tissue, and environment regulate the structural outcome of glucosinolate activation. *Plant Physiol* **147**: 415–428
- West MAL, Kim K, Kliebenstein DJ, Van Leeuwen H, Michelmore RW, Doerge RW, St. Clair DA (2007) Global eQTL mapping reveals the complex genetic architecture of transcript-level variation in Arabidopsis. *Genetics* **175**: 1441–1450

- Wheat CW, Vogel H, Wittstock U, Braby ME, Underwood D, Mitchell-Olds T (2007) The genetic basis of a plant-insect coevolutionary key innovation. *Proc Natl Acad Sci USA* **104**: 20427–20431
- Wittstock U, Agerbirk N, Stauber EJ, Olsen CE, Hippler M, Mitchell-Olds T, Gershenzon J, Vogel H (2004) Successful herbivore attack due to metabolic diversion of a plant chemical defense. *Proc Natl Acad Sci USA* **101**: 4859–4864
- Wittstock U, Burow M (2007) Tipping the scales: specifier proteins in glucosinolate hydrolysis. *IUBMB Life* **59**: 744–751
- Wittstock U, Gershenzon J (2002) Constitutive plant toxins and their role in plant defense. *Curr Opin Plant Biol* **5**: 300–307
- Wittstock U, Halkier BA (2002) Glucosinolate research in the Arabidopsis era. *Trends Plant Sci* **7**: 263–270
- Wittstock U, Kliebenstein DJ, Lambrix V, Reichelt M, Gershenzon J (2003) Glucosinolate hydrolysis and its impact on generalist and specialist insect herbivores. *In* JT Romeo, ed, *Integrative Phytochemistry: From Ethnobotany to Molecular Ecology*, Vol 37. Elsevier, Amsterdam, pp 101–125
- Zabala MD, Grant M, Bones AM, Bennett R, Lim YS, Kissen R, Rossiter JT (2005) Characterisation of recombinant epithiospecifier protein and its over-expression in *Arabidopsis thaliana*. *Phytochemistry* **66**: 859–867
- Zeng ZB, Kao CH, Basten CJ (1999) Estimating the genetic architecture of quantitative traits. *Genet Res* **74**: 279–289
- Zhang Z, Ober JA, Kliebenstein DJ (2006) The gene controlling the quantitative trait locus EPITHIOSPECIFIER MODIFIER1 alters glucosinolate hydrolysis and insect resistance in *Arabidopsis*. *Plant Cell* **18**: 1524–1536
- Zrybko CL, Fukuda EK, Rosen RT (1997) Determination of glucosinolates in domestic and wild mustard by high-performance liquid chromatography with confirmation by electrospray mass spectrometry and photodiode-array detection. *J Chromatogr* **767**: 43–52