The Plant Cell, Vol. 23: 716–729, February 2011, www.plantcell.org

Metabolic Engineering in *Nicotiana benthamiana* Reveals Key Enzyme Functions in *Arabidopsis* Indole Glucosinolate Modification [W]

Marina Pfalz,[a,b,1] Michael Dalgaard Mikkelsen,[c,1,2] Paweł Bednarek,[d] Carl Erik Olsen,[e] Barbara Ann Halkier,[c] and Juergen Kroymann[b,3]

[a] Max Planck Institute for Chemical Ecology, D-07745 Jena, Germany
[b] Laboratoire d'Ecologie, Systématique et Evolution, Université Paris-Sud/Centre National de la Recherche Scientifique, F-91405 Orsay, France
[c] University of Copenhagen, Faculty of Life Sciences, Department of Plant Biology, VKR Research Centre for Pro-Active Plants, DK-1871 Frederiksberg C, Copenhagen, Denmark
[d] Max Planck Institute for Plant Breeding Research, D-50829 Cologne, Germany
[e] University of Copenhagen, Faculty of Life Sciences, Department of Basic Sciences and Environment/Bioorganic Chemistry, DK-1871 Frederiksberg C, Copenhagen, Denmark

Indole glucosinolates, derived from the amino acid Trp, are plant secondary metabolites that mediate numerous biological interactions between cruciferous plants and their natural enemies, such as herbivorous insects, pathogens, and other pests. While the genes and enzymes involved in the *Arabidopsis thaliana* core biosynthetic pathway, leading to indol-3-yl-methyl glucosinolate (I3M), have been identified and characterized, the genes and gene products responsible for modification reactions of the indole ring are largely unknown. Here, we combine the analysis of *Arabidopsis* mutant lines with a bioengineering approach to clarify which genes are involved in the remaining biosynthetic steps in indole glucosinolate modification. We engineered the indole glucosinolate biosynthesis pathway into *Nicotiana benthamiana*, showing that it is possible to produce indole glucosinolates in a noncruciferous plant. Building upon this setup, we demonstrate that all members of a small gene subfamily of cytochrome P450 monooxygenases, *CYP81F*s, are capable of carrying out hydroxylation reactions of the glucosinolate indole ring, leading from I3M to 4-hydroxy-indol-3-yl-methyl and/or 1-hydroxy-indol-3-yl-methyl glucosinolate intermediates, and that these hydroxy intermediates are converted to 4-methoxy-indol-3-yl-methyl and 1-methoxy-indol-3-yl-methyl glucosinolates by either of two family 2 *O*-methyltransferases, termed indole glucosinolate methyltransferase 1 (IGMT1) and IGMT2.

INTRODUCTION

Plants produce an enormous diversity of secondary metabolites (i.e., compounds not directly involved in primary processes of basic growth and development). These include alkaloids, terpenes, phenolic compounds, and others. Tens of thousands of plant secondary metabolites have already been described, but this is only a fraction of what is believed to be present in nature (Wink, 2010). Secondary metabolites often mediate the interaction of plants with their biotic and abiotic environment, determine the food quality of crop plants, and are used as a reservoir of chemical structures for pharmaceutical research. However, the genes and pathways responsible for the generation of structural diversity in plant secondary chemistry are to a large extent unknown. Even for the well-studied glucosinolates, a class of secondary metabolites present in the model plant *Arabidopsis thaliana* and relatives, not all of the genes involved in their biosynthesis have been identified (Sønderby et al., 2010).

Glucosinolates are part of an activated plant defense, the glucosinolate-myrosinase system. Myrosinase-catalyzed glucosinolate hydrolysis leads to the generation of breakdown products with a range of biological activities, including deterrence or increased mortality of nonspecialist insect herbivores and attraction or feeding stimulation of enemies specialized on crucifers. In addition, glucosinolate hydrolysis products determine flavor and taste of Brassicaceous crops and can have beneficial or detrimental effects on the health of humans or livestock upon ingestion.

Arabidopsis produces nearly 40 different aliphatic, aromatic, and indole glucosinolates from three amino acid precursors, Met, Phe, and Trp, respectively (Kliebenstein et al., 2001; Reichelt et al., 2002). The core structure of all glucosinolates is synthesized with a similar set of enzymes (for an overview, see Halkier and Gershenzon, 2006). Structural diversity, however, arises from subsequent modification reactions, such as oxidation, hydroxylation, or methoxylation (Kliebenstein et al., 2005; Kliebenstein,

[1] These authors contributed equally to this work.
[2] Current address: Evolva A/S, Bülowsvej 25, DK-1870 Frederiksberg C, Copenhagen, Denmark.
[3] Address correspondence to juergen.kroymann@u-psud.fr.
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.plantcell.org) is: Juergen Kroymann (juergen.kroymann@u-psud.fr).
[W] Online version contains Web-only data.
www.plantcell.org/cgi/doi/10.1105/tpc.110.081711

2009), and, in the case of Met-derived glucosinolates, also from variation in Met carbon chain extension prior to core structure synthesis (Kroymann et al., 2003; Benderoth et al., 2006, 2009).

One of the last unresolved issues in *Arabidopsis* glucosinolate biosynthesis is which genes are responsible for structural variation in indole glucosinolates. Indole glucosinolates are of interest because their core biosynthetic pathway is intimately linked with biosynthesis of both auxin (Hull et al., 2000) and camalexin, an *Arabidopsis* phytoalexin (Glawischnig et al., 2004). More recently, they have attracted particular attention because modified indole glucosinolates and their hydrolysis products mediate a multitude of plant–enemy interactions. They contribute to plant resistance against aphids (Kim and Jander, 2007; Kim et al., 2008; Pfalz et al., 2009), influence oviposition choice of crucifer specialist insects (de Vos et al., 2008; Sun et al., 2009), and are important for plant innate immunity (Bednarek et al., 2009; Clay et al., 2009).

All investigated *Arabidopsis* accessions contain four different indole glucosinolates: indol-3-yl-methyl (I3M), 4-hydroxy-indol-3-yl-methyl (4OH-I3M), 4-methoxy-indol-3-yl-methyl (4MO-I3M), and 1-methoxy-indol-3-yl-methyl (1MO-I3M) glucosinolate, with I3M being the most abundant indole glucosinolate in leaves and 1MO-I3M the most abundant in roots (Kliebenstein et al., 2001; Brown et al., 2003). We recently identified *CYP81F2* as the gene that underlies *Indole Glucosinolate Modifier 1*, a quantitative trait locus for the accumulation of 4OH-I3M and 4MO-I3M (Pfalz et al., 2007, 2009). *CYP81F2* belongs to a small subfamily of cytochrome P450 monooxygenase genes that consists of four members in *Arabidopsis*: *CYP81F1*, *CYP81F2*, *CYP81F3*, and *CYP81F4*. CYP81F2 catalyzes the conversion of I3M to 4OH-I3M, which in turn is converted to 4MO-I3M by one or several as yet unknown methyltransferase(s). Compared with the wild type, *cyp81f2* null mutants produced significantly less 4OH-I3M and 4MO-I3M, but both metabolites were still detectable in plant tissue (Pfalz et al., 2009), suggesting that other enzymes were also capable of hydroxylating I3M to 4OH-I3M.

We also surmised that gene products related to CYP81F2 were involved in the generation of 1MO-I3M, the other methoxy indole glucosinolate in *Arabidopsis*. However, this reaction is particularly difficult to demonstrate due to the likely instability of the postulated intermediate, 1-hydroxy-indol-3-yl-methyl (1OH-I3M) glucosinolate, which has its hydroxy group attached to the nitrogen of the indole ring. In fact, despite numerous analyses of glucosinolate profiles from a multitude of plants, 1OH-I3M had never been reported (Daxenbichler et al., 1991; Fahey et al., 2001; Clarke, 2010). To overcome these obstacles, we used a *Nicotiana benthamiana* transient expression system that had previously been used successfully to engineer the complete pathway for the production of bioactive benzyl (Geu-Flores et al., 2009) and 4-methylsulfinylbutyl (Mikkelsen et al., 2010) glucosinolate in a noncruciferous plant. We first transformed *N. benthamiana* with the genes for the entire indole glucosinolate

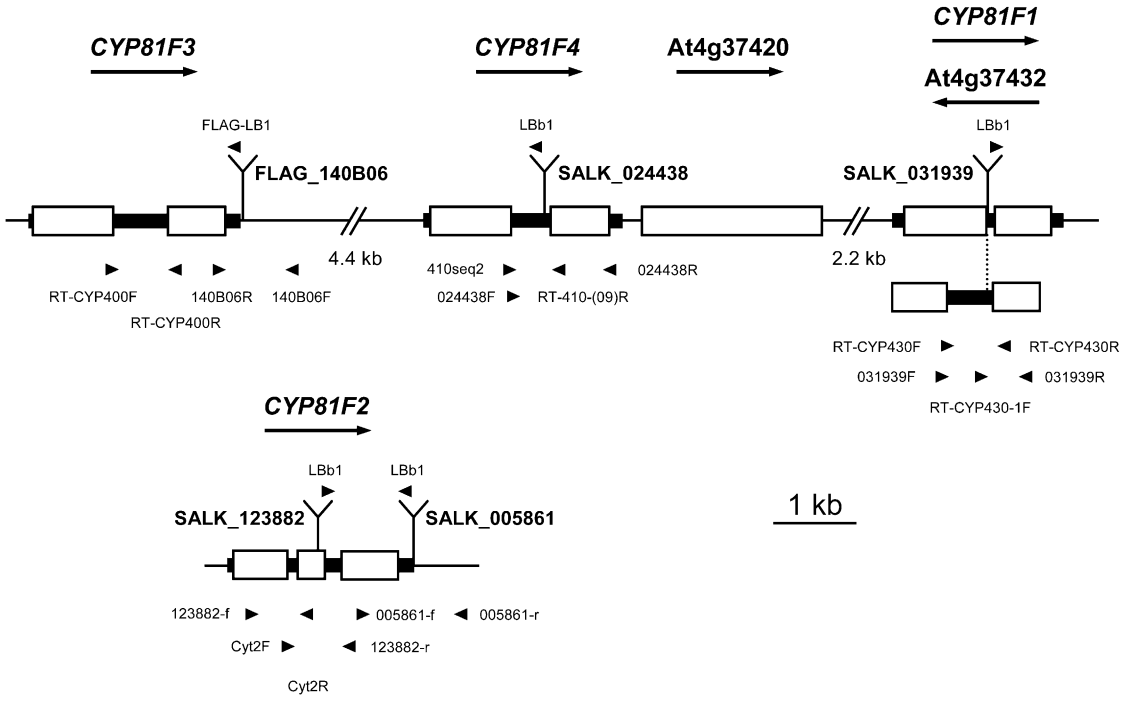

Figure 1. *CYP81F* Genes in *Arabidopsis*.

Shown are coding regions (white bars) and untranslated regions (black bars) of the genes in the *CYP81F* cluster on *Arabidopsis* chromosome 4 (top), with intervening regions shown as thin lines. *CYP81F2* on chromosome 5 is shown for comparison (bottom). Primer positions are indicated as small black triangles. The location of T-DNA insertions in SALK and FLAG lines is indicated by Y-shaped lines. Note that exon-intron structures differ between *CYP81F1* and its antisense gene At4g37432.

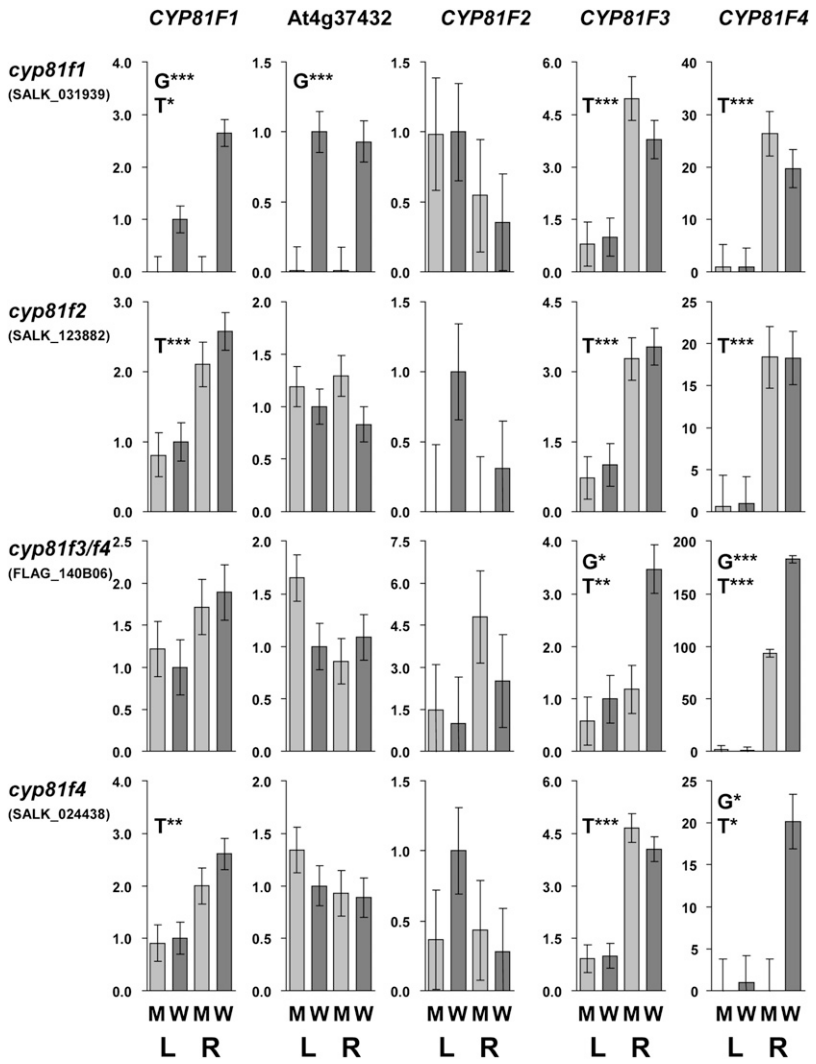

Figure 2. Expression of *CYP81F* Genes in T-DNA Insertion Lines.

Relative transcript levels of *CYP81F* genes and of At4g37432 were assessed in leaves (L) and roots (R) of wild-type (W; dark-gray bars) and mutant (M; light-gray bars) plants in comparison to *eIF4A1* as a control gene. At least three biological replicates were used for each tissue and genotype, with one to three technical replicates per biological replicate. Expression values were set as 1 for wild-type leaves; error bars show SE of the mean. Data were analyzed with analysis of variance: expression = constant + genotype + tissue + genotype*tissue. Statistical significance for differences in transcript levels are indicated for genotype (G) and tissue (T): *P < 0.05, **P < 0.01, and ***P < 0.001.

pathway and then coexpressed *Arabidopsis CYP81F* genes, alone or in combination with candidate *O*-methyltransferase genes, to produce modified indole glucosinolates in this heterologous system.

These analyses show, in combination with assays of T-DNA insertion mutants, that all four members of the *Arabidopsis CYP81F* subfamily have the biochemical capacity to modify I3M, with partially overlapping product spectra. We provide evidence that 1MO-I3M is generated via the same reaction steps as 4MO-I3M and involves a hydroxy intermediate. In addition, we identify a subclade of family 2 plant *O*-methyltransferases implicated in the biosynthesis of modified indole glucosinolates and demonstrate that two of its members, termed indole glucosinolate methyltransferase 1 (IGMT1) and IGMT2, are capable of catalyzing the final step in indole glucosinolate modification (i.e., the transfer of a methyl group to 1OH-I3M and 4OH-I3M to produce 1MO-I3M and 4MO-I3M, respectively).

RESULTS

T-DNA Insertions in *CYP81F* Genes Alter *Arabidopsis* Indole Glucosinolate Profiles

To test our hypothesis that all four *Arabidopsis CYP81F* genes were involved in modification reactions of the indole glucosinolate structure, we first investigated whether T-DNA insertion mutants

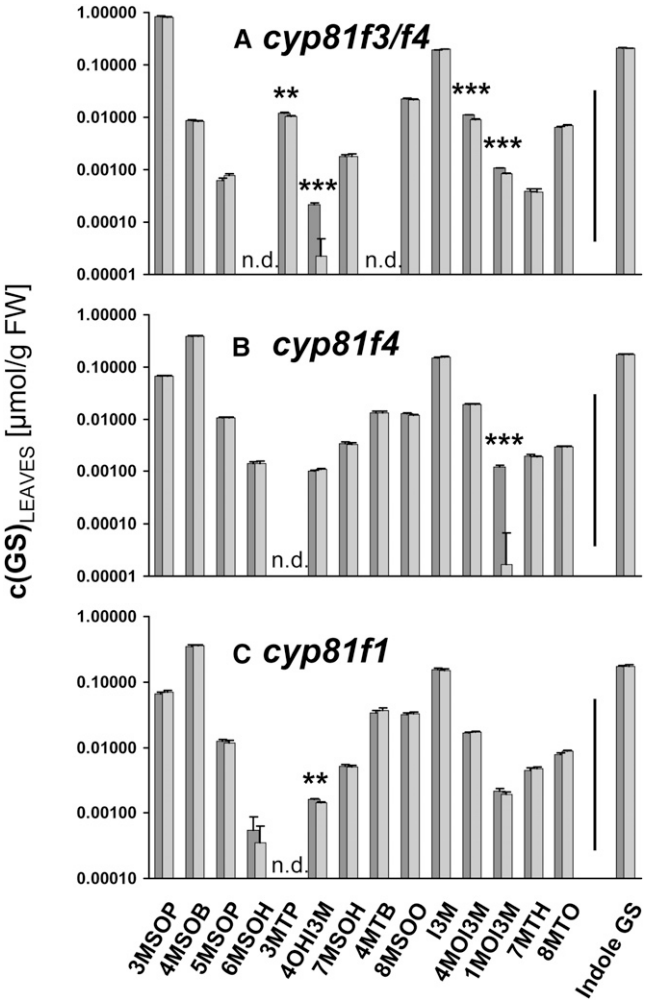

Figure 3. Glucosinolate Profiles in Leaves of *Arabidopsis* T-DNA Insertion Lines.

Glucosinolate levels were measured in leaves of wild-type (dark-gray bars) and mutant plants (light-gray bars) of *cyp81f3/4* **(A)**, *cyp81f4* **(B)**, and *cyp81f1* **(C)**. c(GS) indicates glucosinolate concentration in µmol/g fresh weight (FW). Between 32 and 128 biological replicates were used per genotype and line; error bars show SE of the mean. 3MSOP, 3-methylsulfinyl-propyl; 4MSOB, 4-methylsulfinyl-butyl; 5MSOP, 5-methylsulfinyl-pentyl; 6MSOH, 6-methylsulfinyl-hexyl; 3MTP, 3-methylthio-propyl; 4OHI3M, 4-hydroxy-indol-3-yl-methyl; 7MSOH, 7-methylsulfinyl-heptyl; 4MTB, 4-methylthio-butyl; 8MSOO, 8-methylsulfinyl-octyl; I3M, indol-3-yl-methyl; 4MOI3M, 4-methoxy-indol-3-yl-methyl; 1MOI3M, 1-methoxy-indol-3-yl-methyl; 7MTH, 7-methylthio-heptyl; 8MTO, 8-methylthio-octyl glucosinolate. Bars at the right indicate the total quantity of indole glucosinolates (indole GS). *$P < 0.05$, **$P < 0.01$, and ***$P < 0.001$. n.d., not detected. Note the use of a log scale.

had altered glucosinolate profiles. We obtained T-DNA insertion lines for the *CYP81F* genes from the *Arabidopsis* stock centers (Figure 1). In FLAG_140B06, a line in the Wassilewskija (Ws-4) background (Brunaud et al., 2002), a T-DNA is located between *CYP81F3* and *CYP81F4* downstream of the *CYP81F3* 3′-untranslated region. SALK_031939 has a T-DNA in the intron of *CYP81F1*, SALK_024438 has a T-DNA in the intron of *CYP81F4*, and SALK_123882 has a T-DNA in the second exon of *CYP81F2*. All three SALK lines are in the Columbia (Col) background (Alonso et al., 2003). None of these mutants had any obvious defects in growth or development.

We analyzed all *CYP81F* leaf and root transcript levels in these lines with quantitative RT-PCR (qRT-PCR) to test whether T-DNA insertions caused altered gene expression (Figure 2). As a reference, we chose the gene for the eukaryotic translation initiation factor 4A-1, *eIF4A1* (At3g13920), which displays a very stable expression pattern among different tissues and developmental stages, based on the AtGenExpress database (www.weigelworld.org/resources/microarray/AtGenExpress/). In homozygous SALK_031939 mutants, no *CYP81F1* cDNA was detectable; likewise, we found no *CYP81F4* transcript in homozygous SALK_024438 mutants. For the other *CYP81F*s, we did not detect any statistically significant differences between wild-type and mutant lines. Finally, homozygous FLAG_140B06 mutants had reduced levels of *CYP81F3* ($F = 8.82$, $P = 0.018$, $df = 1$) and *CYP81F4* ($F = 159.27$, $P < 0.001$, $df = 1$) compared with lines with wild-type Ws-4 alleles. Thus, both SALK lines were knockouts for the respective *CYP*s, subsequently designated *cyp81f1* and *cyp81f4*, while FLAG_140B06 was a knockdown for both *CYP81F3* and *CYP81F4* and will be referred to as *cyp81f3/f4*.

We compared glucosinolate levels of T-DNA insertion lines versus lines with wild-type alleles obtained from the same seed material. First, we analyzed leaf glucosinolate profiles for the different T-DNA insertion lines in separate, highly replicated experiments (Figure 3). For *cyp81f3/f4*, we found a highly significant reduction in 4OH-I3M, 4MO-I3M, and 1MO-I3M content (all $n \geq 166$, $df = 1$, $P < 0.001$). In *cyp81f4* leaves, 1MO-I3M was reduced to trace levels ($n \geq 135$, $df = 1$, $P < 0.001$), whereas no other glucosinolates were affected. Finally, *cyp81f1* plants showed some reduction in 4OH-I3M ($n \geq 125$; $df = 1$, $P < 0.01$), but this effect was not very pronounced. Note also that the differences in aliphatic glucosinolate composition (e.g., 3MTP, 4MTP, and 6MSOH) between Ws-4 (Figure 3A) and Col-0 (Figures 3B and 3C) backgrounds result from different alleles at the *MAM* locus, which determines aliphatic glucosinolate carbon chain lengths (Kroymann et al., 2001, 2003; Benderoth et al., 2006, 2009).

Because glucosinolate profiles are tissue specific, we also analyzed glucosinolates in shoots and roots of the same lines. In these experiments, we included the previously identified *CYP81F2* mutant (*cyp81f2*) and wild-type lines (Pfalz et al., 2009) as well as Col-0 for comparison. Shoot data largely confirmed our analysis of single leaves from the previous experiment. 4OH-I3M, 4MO-I3M, and 1MO-I3M were reduced in *cyp81f3/f4*, and 1MO-I3M was nearly undetectable in *cyp81f4*. Likewise, *cyp81f2* had reduced levels of 4OH-I3M and 4MO-I3M, confirming previous findings. Again, *cyp81f1* displayed a somewhat variable phenotype; this time, we found 1MO-I3M levels slightly elevated in mutants. All other glucosinolate levels were indistinguishable between mutants and wild types (Figure 4).

Root phenotypes were, in general, similar to leaves or shoots, with two notable exceptions (Figure 4). First, I3M levels were elevated in *cyp81f2*, *cyp81f3/f4*, and *cyp81f4*. This increase was approximately proportional to the decrease in the levels of

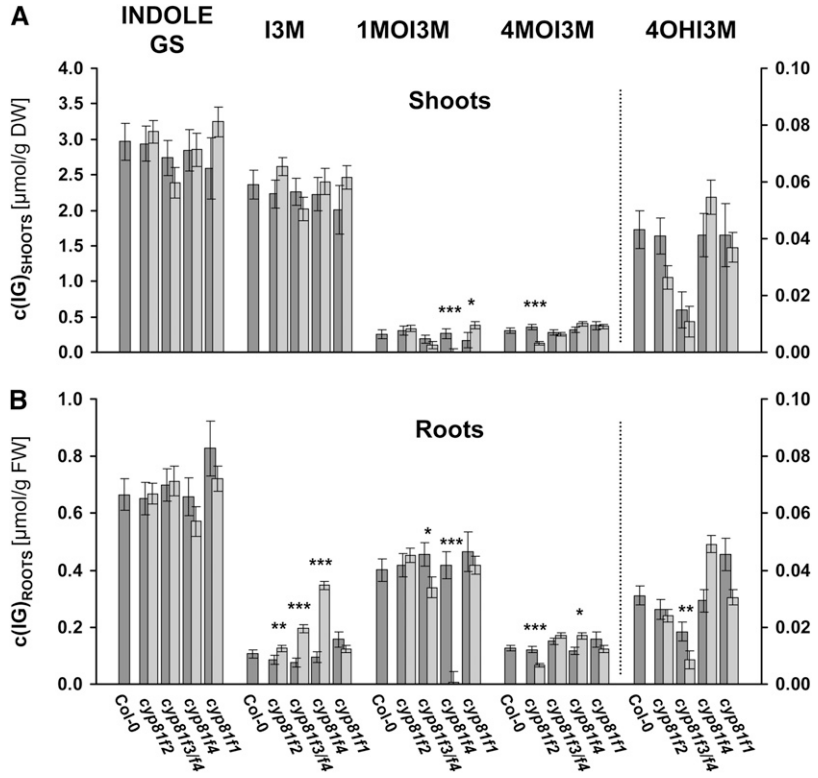

Figure 4. Indole Glucosinolate Profiles in *Arabidopsis* T-DNA Insertion Lines.

(A) Indole glucosinolate profiles in shoots

(B) Indole glucosinolate profiles in roots.

Dark-gray bars indicate the wild type and light-gray bars mutant glucosinolates. *P < 0.05, **P < 0.01, and ***P < 0.001. Between 3 and 21 biological replicates were used per genotype and line; error bars show SE of the mean. Note that we used a different scale for 4OH-I3M.

DW, dry weight; FW, fresh weight.

modified indole glucosinolates, with the strongest impact on I3M in *cyp81f4* and a more moderate effect in *cyp81f3/f4* and *cyp81f2*. Second, in *cyp81f4*, 4OH-I3M and 4MO-I3M levels were also elevated, indicating that a block of 1MO-I3M production in roots resulted in a concomitant increase of 4OH-I3M and 4MO-I3M via accumulation of their precursor I3M.

In summary, the analysis of T-DNA insertion lines indicated that *CYP81F4* was mainly responsible for the production of 1MO-I3M. Furthermore, the comparison of *cyp81f3/f4* and *cyp81f4* mutants suggested that *CYP81F3* was involved in the generation of 4OH-I3M and 4MO-I3M, similar to what was previously reported for *CYP81F2* (Pfalz et al., 2009).

CYP81F3 Catalyzes the Conversion of I3M to 4OH-I3M

Using an expression system based on insect cells, we recently showed that CYP81F2 catalyzed the conversion of intact I3M isolated from seeds of Dyer's woad (*Isatis tinctoria*) to 4OH-I3M (Pfalz et al., 2009). Therefore, we used the same system to express *CYP81F3* and *CYP81F4* heterologously and tested whether the gene products were capable of converting I3M to 4OH-I3M (Figure 5). We included two negative controls in our assays, a buffer control and a heterologously expressed gene from the European cabbage butterfly (*Pieris rapae*). In addition to a major peak of I3M substrate, HPLC revealed a second major peak only in samples with heterologously expressed *CYP81F3* but not in controls or samples with heterologously expressed *CYP81F4*. The retention time of this second peak matched the retention time of 4OH-I3M in an extract of desulfoglucosinolates from *Arabidopsis* roots, and its mass spectrum corresponded to the expected mass of desulfo-4OH-I3M (*m/z* 385 $[M+H]^+$+ *m/z* 223 $[M\text{-Glucose}+H]^+$) in liquid chromatography–mass spectrometry (LC-MS). Thus, based on retention time and LC-MS, we identified this compound as 4OH-I3M.

Transient Coexpression Assays in *N. benthamiana* Reveal That All *Arabidopsis* CYP81Fs Can Modify Indole Glucosinolates

While our enzyme assays based on insect cells clearly showed that both CYP81F2 and CYP81F3 catalyzed the conversion of I3M to 4OH-I3M, we could not demonstrate a role of CYP81F subfamily members in the conversion of I3M to 1OH-I3M with this system. This is likely due to 1OH-I3M instability (Pfalz et al., 2009). As an alternative heterologous expression host, we selected the *N. benthamiana* transient expression system

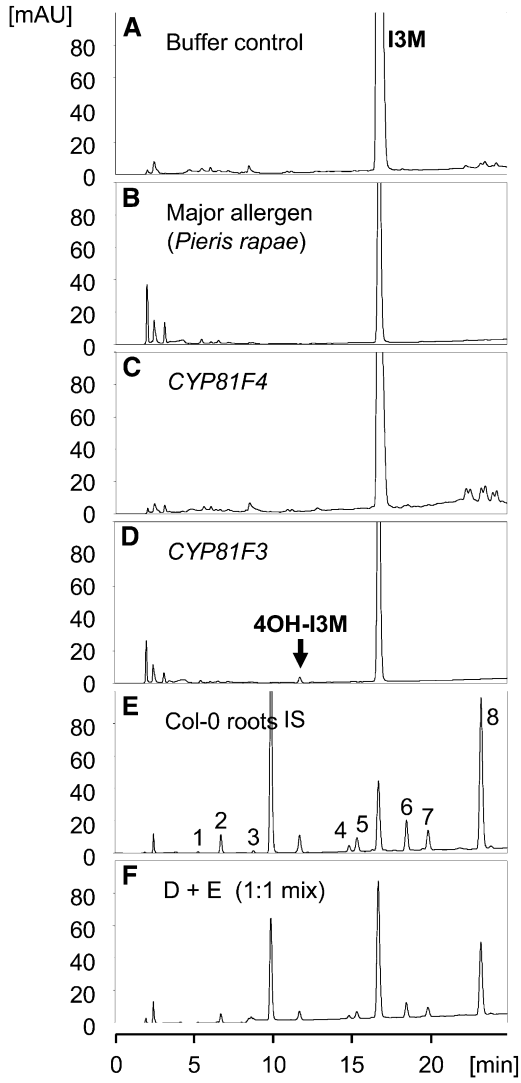

Figure 5. Heterologously Expressed *CYP81F3* Catalyzes the Conversion of I3M to 4OH-I3M.

CYP81F3 and *CYP81F4* were expressed in insect cells and isolated as described in Methods.

(A) and **(B)** Negative controls: a buffer control **(A)** and a heterologously expressed major allergen gene from *P. rapae* **(B)**.

(C) and **(D)** HPLC runs of I3M incubated with heterologously expressed *CYP81F3* **(D)** but not with *CYP81F4* **(C)** have an additional peak corresponding to 4OH-I3M.

(E) and **(F)** Glucosinolates extracted from Col-0 roots **(E)** and a 1:1 mix of **(D)** and **(E)** (shown in **[F]**) are shown for comparison. HPLC runs were done with desulfo-glucosinolates. Additional compounds in **(E)** are 3-methylsulfinyl-propyl (1), 4-methylsulfinylbutyl (2), 5-methylsulfinylpentyl (3), 7-methylsulfinylheptyl (4), 4-methylthiobutyl (5), 8-methylsulfinyloctyl (6), 4-methoxy-indol-3-yl-methyl (7), 1-methoxy-indol-3-yl-methyl (8), and internal standard (IS = *p*-hydroxy-benzyl) desulfo-glucosinolate. mAU, milliabsorption units at 226-nm wavelength.

(Geu-Flores et al., 2009; Mikkelsen et al., 2010), which enabled coupling of the expression of *CYP81F* genes with *O*-methyltransferase (*O-MT*) genes, assuming that we could engineer the production of indole glucosinolate. We expected that these *O*-MTs would carry out the subsequent reaction step toward the more stable 1MO-I3M and toward 4MO-I3M, respectively.

To delimit the large number of *O-MT* genes we searched the ATTED-II database (Obayashi et al., 2009) for genes coexpressed with *CYP81F2*. This tool had previously been successfully used to identify structural genes and transcription factors involved in glucosinolate biosynthesis, based on coexpression with known genes from the glucosinolate biosynthesis pathways (Hirai et al., 2007; Sawada et al., 2009; Albinsky et al., 2010). In the *CYP81F2* coexpression network, two genes encoding *O*-MTs, At1g21100 and At1g21120, were direct neighbors of *CYP81F2* with correlations of $r = 0.618$ and $r = 0.742$, respectively. A third direct neighbor of *CYP81F2* in this network was *MYB51*, encoding an R2R3-MYB transcription factor that had recently been shown to regulate indole glucosinolate biosynthesis in *Arabidopsis* (Gigolashvili et al., 2007). We tested the correlation of gene expression between *CYP81F2* and both methyltransferase genes with a publicly available set of 1436 expression experiments with Affymetrix *Arabidopsis* ATH1 microarrays (www.Arabidopsis.org) and found a statistically significant correlation for the expression of both gene pairs, *CYP81F2* and At1g21100 ($n = 1436$, $r = 0.58$, $P < 0.00001$), and *CYP81F2* and At1g21120 ($n = 1436$, $r = 0.47$, $P < 0.00001$), respectively.

Based on the same data set, *CYP81F3* and *CYP81F4* (but not *CYP81F1*) displayed a statistically significant correlation in expression with At1g21100 and/or At1g21120 and also with three other related *O-MT* genes (At1g21110, At1g21130, and At1g76790). Furthermore, correlations in expression between *CYP81F*s, with the exception of *CYP81F1*, and the five *O-MT*s were also significant in microarray subsets specific for development, hormone treatments, abiotic stress, or pathogens (Schmid et al., 2005). These data are shown in Supplemental Figures 1 to 3 online.

Nicotiana and all other solanaceous species naturally lack the entire glucosinolate-myrosinase complex. In this system, we introduced expression constructs for the complete I3M pathway (Figure 6). *Agrobacterium tumefaciens* strains containing constructs harboring *CYP79B2*, *CYP83B1*, *GSTF9*, *GGP1*, *SUR1*, *AtST5a*, and *UGT74B1* were coinfiltrated into *N. benthamiana* leaves along with the P19 suppressor strain (Voinnet et al., 2003). The functions of the individual genes are described in Methods.

After 7 d of incubation, we extracted glucosinolates and analyzed their desulfo-glucosinolate derivatives by LC-MS. This procedure identified a single peak of *m/z* 391 $[M+Na]^+$ corresponding to I3M (Figure 7; see Supplemental Figure 4 online). The concentration of I3M was ~0.2 nmol/mg fresh weight, which is comparable to what has been reported for benzyl glucosinolate produced in *N. benthamiana* (Geu-Flores et al., 2009). Then we coexpressed *CYP81F* and/or *O-MT* genes with the genes of the I3M core pathway. We extracted glucosinolates and subjected them to LC-MS. We focused on extracted ion chromatograms for molecular weights corresponding to I3M (*m/z* 391) and its hydroxylated (OH-I3M; *m/z* 407), methoxylated (MO-I3M; *m/z* 421), hydroxylated and methoxylated

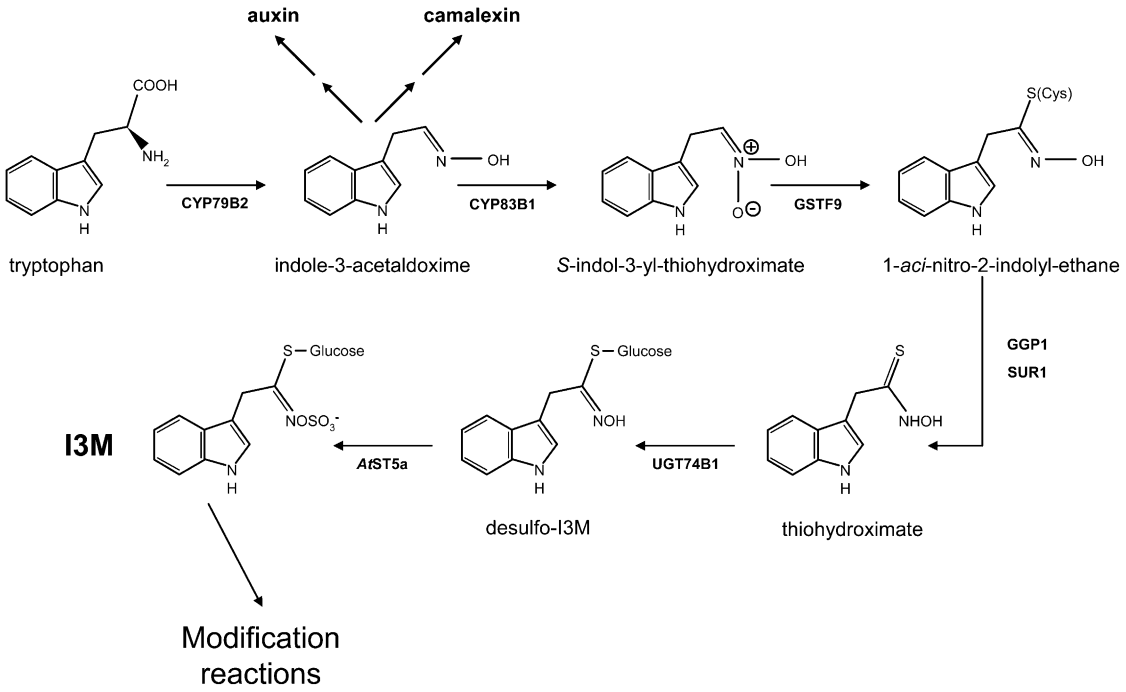

Figure 6. Reconstitution of the *Arabidopsis* Indole Glucosinolate Core Pathway in *Nicotiana*.

Shown are intermediates starting from Trp along with the gene products used for transient reconstitution of the indole glucosinolate core pathway in *N. benthamiana*. I3M, the product of the indole glucosinolate core pathway, is the substrate in subsequent modification reactions, shown in Figure 9.

(OH-MO-I3M; *m/z* 437), and/or double methoxylated (2xxMO-I3M; *m/z* 451) forms, respectively. We further analyzed these compounds with LC-MS/MS to verify the presence of *S*-glucose, which indicates the presence of a glucosinolate.

First, we assayed the *CYP81F* genes without *Arabidopsis O*-MTs and identified products based on retention time, UV spectra, and LC-MS/MS. Compared with controls, we detected a novel peak with *m/z* 407 in assays with *CYP81F1*, *CYP81F2*, and *CYP81F3* but not with *CYP81F4*. This peak eluted at the same time as an authentic standard of 4OH-I3M (Figure 7; see Supplemental Figure 4 online). Next, we tested the *CYP81F* genes together with individual *O-MT* genes, either with At1g21100 or with At1g21120. At standard resolution, we identified a novel product with *m/z* 421 that eluted at the same time as 4MO-I3M isolated from seedlings of *Arabidopsis* Col-0 (Figure 8; see Supplemental Figures 5 and 6 online). This product was present in assays with *CYP81F1*, *CYP81F2*, and *CYP81F3* but absent in assays with *CYP81F4*. The product spectrum was the same with each individual *O-MT*, At1g21100 or At1g21120 (see Supplemental Figure 7 online). At higher resolution, we also detected a second novel product with *m/z* 421 that eluted at the same time as 1MO-I3M from *Arabidopsis* seedlings. This product was present in all assays with *CYP81F* genes, including *CYP81F4* (Figure 8; see Supplemental Figures 5 and 6 online). Products with *m/z* 407 or *m/z* 421 that eluted at the same time as 4OH-I3M, 4MO-I3M, or 1MO-I3M were absent when *O-MT* genes were assayed without *CYP81F* genes (Figure 8; see Supplemental Figure 7 online).

Thus, heterologous expression of *CYP81F*s alone or in combination with either of two *O-MT* genes showed that all four *CYP81F* genes had the biochemical capacity to convert I3M to 1OH-I3M and, except *CYP81F4*, also to 4OH-I3M. Furthermore, both tested *O*-MTs accepted the hydroxy intermediates as substrates for the generation of 1MO-I3M and 4MO-I3M, respectively.

DISCUSSION

The Indole Glucosinolate Modification Pathways

All four *Arabidopsis CYP81F* gene products had the capacity to modify the indole glucosinolate structure. CYP81F1, CYP81F2, and CYP81F3, but not CYP81F4, catalyzed the conversion of I3M to 4OH-I3M, and all four CYP81Fs converted I3M to 1OH-I3M (Figure 9). We could infer this latter activity because 1MO-I3M appeared in our assays only when *CYP81F* and *Arabidopsis O-MT* genes were present but not when *O-MT*s were expressed without *CYP81F* genes.

By contrast, our analyses of T-DNA insertion lines investigated the impact of *cyp81f* mutations in the presence of the remaining intact *CYP81F* genes in *Arabidopsis*. Glucosinolate profiles showed that *CYP81F4* was mainly responsible for the generation of 1MO-I3M because this metabolite was virtually absent in the investigated root, shoot, and leaf tissues of *cyp81f4* mutants. This defect was not compensated by *CYP81F1*, *2*, or *3*, even

though their gene products had the biochemical capacity to convert I3M to 1OH-I3M in the *Nicotiana* expression system. Instead, levels of all other indole glucosinolates were increased. This effect was particularly pronounced in root tissue where 1MO-I3M constituted the most abundant indole glucosinolate (Figure 4) and where *CYP81F4* was strongly expressed (Figure 2; see Supplemental Figure 3 online). Since *CYP81F1*, *2*, and *3* transcript levels were unchanged in *cyp81f4* mutants compared with *CYP81F4* wild type (Figure 2), this effect was likely due to shifts in reaction equilibria: the substrate of all CYP81Fs, I3M, accumulated due to the lack of functional CYP81F4 and additional I3M was partially converted into 4OH-I3M and then 4MO-I3M. However, total indole glucosinolate levels remained unchanged in roots and shoots in all investigated mutants, indicating that *CYP81F* genes did not control the absolute quantity of indole glucosinolates.

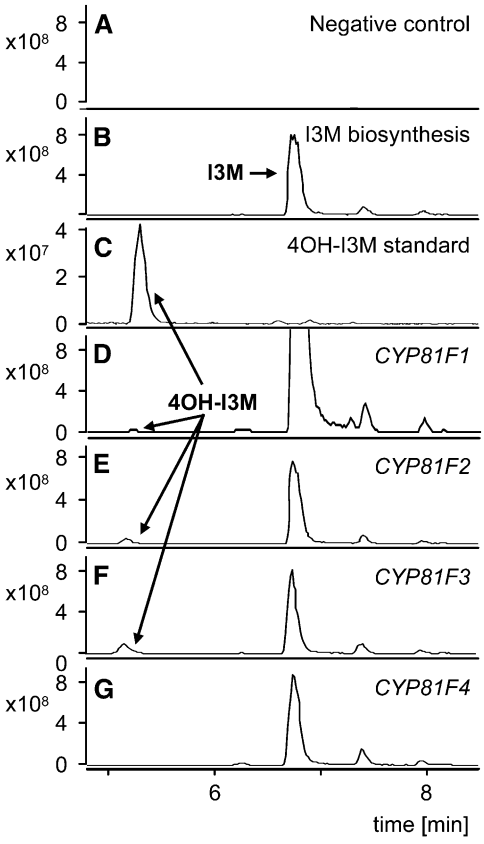

Figure 7. Engineering of *CYP81F* Genes in *N. benthamiana*.

(A) Negative control consisting of infiltrated *Agrobacterium* without any expression constucts.

(B) LC-MS traces of plants transformed with genes for the indole glucosinolate core pathway revealed a novel compound, I3M.

(C) 4OH-I3M standard.

(D) to **(G)** Additional transformation with *CYP81F1* **(D)**, *CYP81F2* **(E)**, or *CYP81F3* **(F)** but not with *CYP81F4* **(G)** led to the generation of an additional peak corresponding to 4OH-I3M, based on a comparison of elution time with an authentic standard, on UV spectra, and on LC-MS data. The *y* axis indicates relative ion abundance.

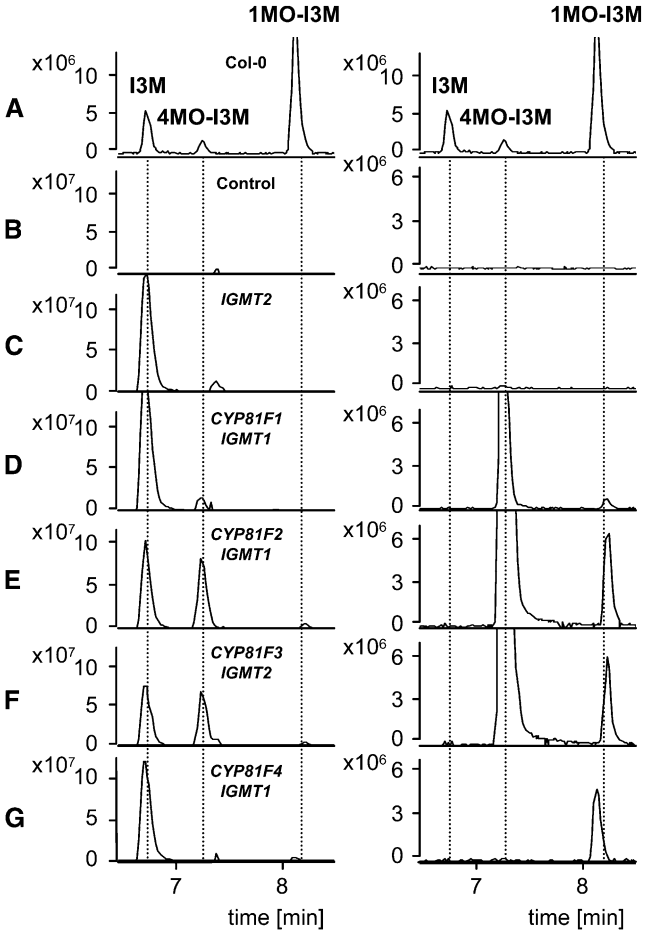

Figure 8. Engineering of *CYP81F* and *O-MT* Genes in *N. benthamiana*.

(A) and **(B)** Glucosinolate profiles from Col-0 seedlings **(A)** and from a negative control consisting of *Agrobacterium* without any expression constructs **(B)** are shown for comparison.

(C) *O-MT* genes alone have no impact on the glucosinolate profile of *Nicotiana* transformed with the indole glucosinolate core pathway genes.

(D) to **(G)** Coexpression of *O-MT* and *CYP81F* genes leads to the formation of methoxy indole glucosinolates. Coexpression of *O-MT*s with *CYP81F1* **(D)**, *CYP81F2* **(E)**, and *CYP81F3* **(F)** but not with *CYP81F4* **(G)** leads to the generation of 4MO-I3M, and all *CYP81F*s produce 1MO-I3M when coexpressed with *IGMT1* or *2*. Note that the elution time of 1MO-I3M slightly decreases when preceeded by large amounts of 4MO-I3M. The *y* axis indicates relative ion abundance.

Shown are glucosinolate profiles at standard (left) and enhanced (right) resolution. The left panels focus on extracted ion chromatograms for molecular weights corresponding to I3M (*m/z* 391) and its hydroxylated (OH-I3M; *m/z* 407), methoxylated (MO-I3M; *m/z* 421), hydroxylated and methoxylated (OH-MO-I3M; *m/z* 437), and/or double methoxylated (2xxMO-I3M; *m/z* 451) forms, respectively. Thus, these panels are filtered to reveal the substrate and all theoretically possible reaction products of CYP81Fs coupled with individual *O*-MTs. The right panels focus on compounds with *m/z* 421 (i.e., the expected methoxylated reaction products). All other peaks not corresponding to this molecular weight are therefore blanked out, including I3M and some minor contaminants present also in the control. Compounds were identified based on comparison with authentic standards, UV spectra, and LC-MS data.

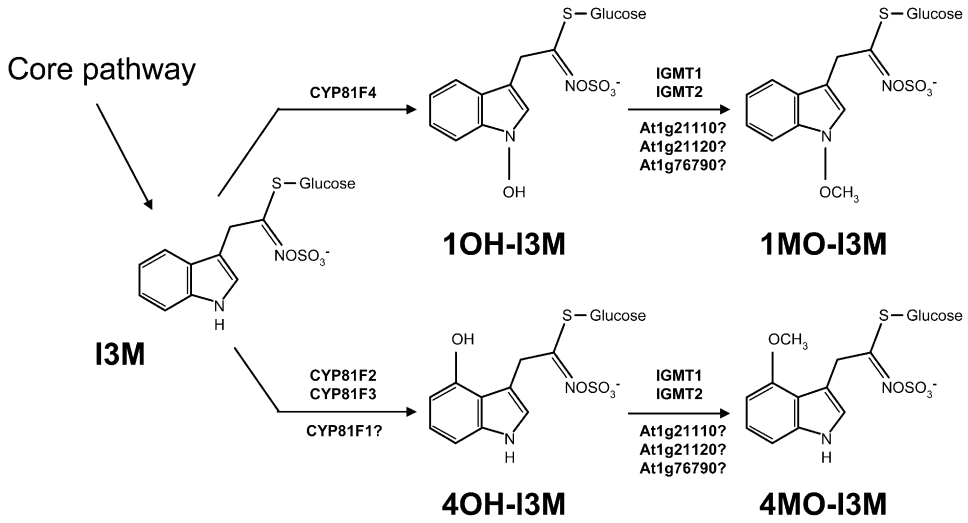

Figure 9. The Indole Glucosinolate Modification Pathway in *Arabidopsis*.

Enzyme functions inferred from a combination of *Arabidopsis* T-DNA insertions and biochemical assays are shown above arrows, and further gene products that may catalyze the same reaction are shown below arrows. I3M is the product of the indole glucosinolate core pathway and acts as the substrate for hydroxylation reactions performed by members of the CYP81F subfamily of cytochrome P450s. Hydroxy intermediates are further converted to 1MO-I3M and 4MO-I3M, respectively, by IGMTs, which belong to family 2 plant *O*-MTs.

Previous work had shown that CYP81F2 catalyzed the conversion of I3M to 4OH-I3M, which is in turn the substrate for production of 4MO-I3M (Pfalz et al., 2009). Nonetheless, *cyp81f2* mutants still accumulated considerable amounts of 4OH-I3M and 4MO-I3M (Bednarek et al., 2009; Clay et al., 2009; Pfalz et al., 2009), suggesting that at least one other gene was able to catalyze the same reaction. In this work, we show with heterologously expressed *CYP81F3* in insect cell lines (Figure 5) and in *Nicotiana* (Figure 7) that CYP81F3 can catalyze the hydroxylation of I3M to 4OH-I3M. Moreover, leaves of *cyp81f3/f4* knockdown mutants had reduced levels of 4OH-I3M, 4MO-I3M, and 1MO-I3M (Figure 3). The comparison with *cyp81f4* knockout lines indicated that the reduction of 1MO-I3M levels in *cyp81f3/f4* was attributable to the knockdown of *CYP81F4*. Hence, the knockdown of *CYP81F3* accounted for reduced levels of 4OH-I3M and 4MO-I3M in *cyp81f3/f4* mutants. Together, biochemical assays and mutant analyses thus indicate that CYP81F3 catalyzes in planta the same reaction as CYP81F2, the conversion of I3M to 4OH-I3M. It remains to be seen whether a *cyp81f2*/*cyp81f3* double knockout completely abolishes 4OH-I3M and 4MO-I3M synthesis or whether additional genes, possibly *CYP81F1*, contribute to the generation of these metabolites in planta. Crosses between FLAG_140B06 (*cyp81f3/f4* double knockdown in Ws-4 background) and SALK_005861 or SALK_123882 (both *cyp81f2* knockouts in Col-0 background; Pfalz et al., 2009) were inconclusive because major glucosinolate quantitative trait locus segregated in the progeny from these crosses (data not shown).

The biological role of *CYP81F1* remained somewhat mysterious. While our biochemical assays clearly indicated that heterologously expressed *CYP81F1* could catalyze the formation of both 1OH-I3M and 4OH-I3M from I3M, *cyp81f1* mutants did not show a consistent glucosinolate phenotype. This absence of a glucosinolate phenotype may be caused by a natural antisense gene, At4g37432, which largely overlaps with *CYP81F1* (Figure 1). We confirmed with qRT-PCR that the antisense transcript was present in the wild type but not in *cyp81f1* mutants. If this antisense transcript abolished *CYP81F1* translation in the wild type, this would explain why *cyp81f1* mutant glucosinolate profiles did not differ from wild-type plants in the tissues that we investigated.

Both At1g21100 and At1g21120 gene products catalyzed the transfer of methyl groups to hydroxy indole glucosinolates, leading from 4OH-I3M to 4MO-I3M and from 1OH-I3M to 1MO-I3M, respectively (Figure 9). Plant species harbor large numbers of *O-MT* genes (Noel et al., 2003; D'Auria, 2006; Lam et al., 2007). The exact number of *O-MT*s in *Nicotiana* is unknown, but multiple *O-MT* genes have been identified (Pinçon et al., 2001; Lam et al., 2007; Hippauf et al., 2010). However, endogenous *Nicotiana* *O*-MTs were not capable of converting I3M to either 1OH-I3M or 4OH-I3M. We can therefore conclude that the *Arabidopsis O-MT* genes that we coexpressed with *CYP81F* genes are fairly specific for the biosynthesis of modified indole glucosinolates, and we thus termed them *IGMT1* and *IGMT2*. *IGMT1* and *2* form, together with At1g21110 and At21130, a small gene cluster of *O-MT* genes on *Arabidopsis* chromosome 1. These four *O*-MTs are 95 to 98% identical at the amino acid level and share ~70% sequence identity with a fifth *O*-MT, At1g76790, located near the bottom of *Arabidopsis* chromosome I. All five gene products belong to family 2 of plant *O*-MTs but are only distantly related to other characterized plant *O*-MTs. The best match to an *O*-MT whose substrate specificity has been determined (Lam et al., 2007) is an *O*-MT from *Rosa chinensis* involved in the biosynthesis of floral scent compounds (Wu et al., 2004), with an amino acid identity of only around 50%. Thus, the two IGMTs identified in this work, together with

At1g21110, At2g21130, and At76790, form a group of *O*-MTs that is only distantly related to other plant *O*-MTs. Whether these latter genes also function in indole glucosinolate modification remains to be seen.

Why four *CYP81F* genes with partially or entirely overlapping product spectra are present in *Arabidopsis* also remains to be investigated. Microarray-based transcript profiling data from AtGenExpress, a collection and analysis tool for gene expression data (Schmid et al., 2005), indicate that *CYP81F*s may be expressed in an organ- or tissue-specific manner (see Supplemental Figure 3 online). These data show, for example, that *CYP81F4* is preferentially (but not exclusively) expressed in roots, consistent with a high accumulation of 1MO-I3M in this tissue. However, *CYP81F4* transcript was also present in wild-type leaves (Figure 2), and *cyp81f4* mutants displayed a strong reduction in 1MO-I3M levels not only in roots but also in leaves (Figures 3 and 4). Another possibility is that the four *CYP81F* genes serve in distinct but parallel pathways. In this context, it is worth noting that *CYP81F4* is coexpressed with *PYK10* (At3g09260), a root- and hypocotyl-specific β-glucosidase gene (Nitz et al., 2001; Sherameti et al., 2008). The *PYK10* gene product has ~50% identity to PEN2 at the amino acid level, and PEN2 was recently identified as an atypical myrosinase responsible for hydrolyzing *CYP81F2*-dependent 4MO-I3M in plant innate immunity reactions (Bednarek et al., 2009; Clay et al., 2009). However, heterologously expressed *PYK10* did not accept sinigrin, an aliphatic glucosinolate, as a substrate (Ahn et al., 2010), but this does not preclude a potential activity with indole glucosinolates.

Transient Expression in *N. benthamiana* as a Platform for Investigating Biochemical Pathways

To determine the function of gene family members is a challenging task, in particular when some of its members form a cluster of tightly linked genes. Heterologous expression of *CYP81F* genes in insect cell lines was only partially successful in characterizing gene product function, but transient expression in *Nicotiana* made it possible to demonstrate the biochemical capacity of all four members of the *Arabidopsis CYP81F* gene family. Of course, these data have to be interpreted cautiously because expression conditions in *Nicotiana* do not completely reflect the native cellular environment in *Arabidopsis* or other plants. Expression of glucosinolate biosynthesis genes in *Nicotiana* was driven by strong promoters. This may lead to an artificial stoichiometry of the gene product and metabolite ratios and is a potential explanation for the generation of 1MO-I3M in assays with CYP81F1, 2, or 3 coupled with *O*-MTs. These reactions, according to mutant glucosinolate profiles, apparently do not occur in *Arabidopsis*. It is also not completely clear whether the *O-MT*s chosen for coexpression in *Nicotiana* are expressed in the same cells and tissues as the *CYP81F*s in *Arabidopsis*, although in silico analyses (see Supplemental Figures 1 to 3 online) strongly suggest that this is the case. Finally, it remains possible that other as yet unidentified factors influence indole glucosinolate biosynthesis in *Arabidopsis*.

This *Nicotiana* expression system has previously been used to successfully engineer the pathways for Phe-derived benzyl glucosinolates (Geu-Flores et al., 2009) and for Met-derived aliphatic glucosinolates (Mikkelsen et al., 2010). This work shows the same for indole glucosinolate biosynthesis. Thus, the core pathways for all three classes of *Arabidopsis* glucosinolates can be reconstituted in *Nicotiana*. However, this work also shows that it is possible to further manipulate these pathways and include modification reactions for the generation of glucosinolates that bear various structural substitutions. This demonstrates that transient expression in *N. benthamiana* is a useful platform for validation of genes in biosynthetic pathways and, moreover, that this system has potential for the generation of plant secondary compounds tailored for specific needs.

METHODS

Plant Growth Conditions

Arabidopsis thaliana T-DNA insertion lines were obtained from the Nottingham Arabidopsis Stock Centre (NASC) or Institut National de la Recherche Agronomique, Versailles. After sowing into damp potting medium, seeds were cold treated at 6°C for 3 to 4 d. Seeds were germinated in the light, and seedlings were either transferred to 1:3 vermiculite/potting soil mix with 20 mL slow release fertilizer (Osmocote) per flat or to quartz sand (1:1 mix of 0.7- to 1.2-mm and 3- to 4-mm grain sizes) where a fertilization with 300 mL 0.1% (w/v) 15-10-15/N-P-K fertilizer, supplied with micronutirents (Planta Düngemittel), was conducted after 18 d of plant culture. Plant were grown in an environmentally controlled growth chamber equipped with NH 360 FLX Sunlux ACE bulbs with 200 μmol s^{-1} m^{-2} as a light source, with 11.5-h-day/12.5-h-night cycles at 22°C and 60% relative humidity (day) and 16°C and 80% relative humidity (night). Plants grown in soil were used at 4 weeks age for the analysis of leaf glucosinolates, and for RNA extraction prior to qRT-PCR, plants were grown in soil at 5 weeks age for root and shoot glucosinolate profiles. *Nicotiana benthamiana* plants were grown in the greenhouse with 16-h-day/8-h-night cycles at 28°C (day) and 25°C (night).

DNA Extraction and Genotyping

Genomic DNA was isolated from one to two freeze-dried *Arabidopsis* leaves as described by Kroymann et al. (2001). Seed material for T-DNA insertion lines provided by NASC was segregating for the T-DNA insertions. Lines homozygous for either mutant or wild-type alleles were identified using the PCR-based markers listed in Supplemental Table 1 online and shown in Figure 1. For example, primers 031939F + R produced a PCR product when wild-type alleles were present in progeny from SALK_031939 and primers LBb1 + 031939R when a plant carried a mutant allele. In general, PCR was performed with ~30 ng DNA, 2.3 μL 10× PCR buffer (Qiagen), 4 nmol of each deoxynucleotide triphosphate, 1.25 pmol of each of both primers, 70 nmol $MgCl_2$, and 0.15 units of *Taq* DNA polymerase (Qiagen) in a 23-μL volume on an Applied Biosystems 9700 Thermocycler, with 94°C for 2 min, followed by 38 cycles of 94°C for 30 s, 50°C or 55°C for 15 s, and 72°C for 30 s, and a final extension of 72°C for 2 min. The number of T-DNA insertions in all three mutant lines was examined by nonradioactive DNA gel blot hybridization (see Supplemental Figure 8 online) using the Amersham ECL direct nucleic acid labeling and detection system (GE Healthcare). Per plant, 2 μg of total genomic DNA were digested with either *Bam*HI (FLAG_140B06 and SALK_031939) or with *Hin*dIII (SALK_024438). After gel electrophoresis and blotting, samples were hybridized with an ~300-bp fragment obtained from the T-DNA left border region using primer pairs T-DNA_L1F and T-DNA_L1R or FLAG-F and FLAG-R, respectively.

RNA Extraction and Quantitative Real-Time PCR

After harvest, roots were briefly washed to remove attached soil, and roots and leaves were quick frozen in liquid N_2. Plant tissue was ground to a fine powder, and total RNA was extracted with TRIsure reagent (Bioline) following the protocol provided by the manufacturer. After a treatment with DNase (Turbo DNase; Ambion), a second purification was performed with RNeasy MinElute columns (Qiagen) to remove DNase and any contaminating polysaccharides and proteins. RNA Nano chips (Agilent Technologies) were used to verify RNA integrity on an Agilent 2100 Bioanalyzer, and RNA quantity was determined with a Nanodrop ND-1000 spectrophotometer (Nanodrop Technolgies) or a BioPhotometer 6131 (Eppendorf). qRT-PCR was performed with 500 ng of DNA-free total RNA converted into single-stranded cDNA using a mix of random and oligo(dT_{20}) primers according to the ABgene protocol. Gene-specific primers for qRT-PCR are listed in Supplemental Table 1 online. Primer design, qRT-PCR conditions, dissociation curve analysis, tests of dynamic range, and quantification were performed as described by Freitak et al. (2007). qRT-PCR was conducted on an Applied Biosystems StepOnePlus System using SYBR Green Reagents using the comparative C_T ($\Delta\Delta C_T$) method with *eIF4A1* as a reference gene. The temperature profile was 95°C for 15 min, followed by 40 cycles of 95°C for 15 s, 55.5°C for 30 s, and 72°C for 30 s, followed by a melt curve stage with 95°C for 30 s, 60°C for 30 s, and a subsequent incremental increase of +0.5°C every 15 s until a final temperature of 95°C was reached. At least three biological replicates were used per tissue and genotype, and one to three technical replicates were used per biological replicate.

Glucosinolate Analysis of *Arabidopsis* Tissues

For analysis of leaf glucosinolate composition, 100 mg leaf material was used from plants grown in soil. Plants grown in sand were used for analysis of shoot and root glucosinolates. Here, complete shoots and roots were harvested separately. Roots were briefly rinsed to remove attached sand and then quickly dried with a paper towel. Shoots and roots were weighted to determine fresh weight. Tissue was quick frozen in liquid N_2 and lyophilized to dryness. Prior to glucosinolate extraction, samples were homogenized with 2.3-mm ball bearings in a paint shaker and incubated for 15 min in 1 mL 80% (v/v) aqueous methanol. *p*-Hydroxybenzyl-glucosinolate (0.05 mM) was added to each sample as an internal standard. Glucosinolates were extracted in a 96-well format. After 10 min of centrifugation at 2000*g*, the supernatants were loaded onto Sephadex DEAE 25A columns, and four washing steps were performed (1× 0.5 mL 80% [v/v] aqueous methanol, 2× 1 mL deionized water, and 1× 0.5 mL 0.02 M MES buffer, pH 5.2). Glucosinolates were incubated overnight with *Helix pomatia* sulfatase, and desulfo-glucosinolates were eluted with 0.5 mL deionized water. HPLC separation, desulfo-glucosinolate identification, and quantification were performed according to Kroymann et al. (2001). Response factors from Brown et al. (2003) were used to calculate molar glucosinolate concentrations.

Heterologous Expression in *Sf*9 cells

Total RNA from *Arabidopsis* Col-0 was extracted with TRIzol (Invitrogen) and reverse transcribed with SuperScript III reverse transcriptase (Invitrogen) following the manufacturer's instructions. *CYP81F3* and *CYP81F4* cDNA were amplified with AccuPrime Taq polymerase (Invitrogen) with primer pairs 400_Ins-Exp-KOZ-F and 400_Ins-R, and with 410_Ins-Exp-KOZ-F and 410_Ins-R, respectively (see Supplemental Table 1 online). Primer design followed the rationale presented by Pfalz et al. (2009) so that the PCR products contained a Kozak translation initiation sequence (Kozak, 1987, 1990, 1991) in addition to the start codon, while the native stop codon was omitted to include a C-terminal vector-encoded peptide for detection with a V5 antibody. PCR was conducted on a PE Applied Biosystems 9700 thermal cycler, with 94°C for 1 min, 30 cycles at 94°C for 15 s, 55°C for 30 s, and 72°C for 1.45 min, followed by a final extension for 5 min at 72°C. PCR products were gel purified, cloned into a PIB/V5-His TOPO vector (Invitrogen), and transformed into *Escherichia coli* TOP-10 cells (Invitrogen). Following isolation with the HiPure Plasmid Filter Midiprep kit (Invitrogen), plasmids were sequenced to verify correct cDNA sequence, reading frame, and cloning direction.

*Spodoptera frugiperda Sf*9 cells (Invitrogen) were grown in culture dishes at 27°C in *Sf*-900 II SFM with 50 μg/mL Gentamycin (both Gibco). Per dish, 660 μL SFM was mixed with 66 μL Insect GeneJuice transfection reagent (Novagen). This mixture was combined with 660 μL SFM containing 12 μg of expression construct and left for 15 min at ambient temperature before adding to *Sf*9 cells. SFM was replaced once after 4 h incubation at 27°C.

After 48 h at 27°C, cells were harvested with their culture medium, centrifuged for 10 min with 500*g* at 4°C, and washed twice with ice-cold 1× PBS. Cells were resuspended in 2 mL hypotonic buffer (20 mM Tris-HCl, pH 7.5, 5 mM EDTA, 1 mM DTT, and 1× protease inhibitor cocktail [Pierce]) and placed on ice for 20 min. Afterwards, cells were homogenized and mixed with an equal volume of sucrose buffer (20 mM Tris-HCl, pH 7.5, 5 mM EDTA, 1 mM DTT, 500 mM sucrose, and protease inhibitors) and centrifuged for 10 min with 1200 g at 4°C. The supernatant was stored on ice. Homogenization was repeated once with the pellet. Pooled supernatants were centrifuged for 15 min with 10,000*g* at 4°C, and the resulting supernatant was centrifuged for 1 h with 100,000*g* at 4°C. The pellet from this step was resuspended in 0.5 mL 100 mM phosphate buffer, pH 7.4, with 20% (v/v) glycerol and protease inhibitors. Subsequently, heterologous expression of *CYP81F* genes was confirmed with SDS-PAGE and protein gel blotting.

Enzyme Assays

Activity of heterologously expressed proteins encoded by *CYP81F* genes was measured in microsomes isolated from *Sf*9 insect cells. In parallel, assays were conducted with microsomes isolated from *Sf*9 cells expressing a major allergen gene from the European cabbage butterfly (*Pieris rapae*) as one negative control (Fischer et al., 2008) and phosphate buffer instead of microsomes as a second negative control. Bio-Rad DC protein assays were used to determine protein concentration in microsomes expressing *CYP81F* genes, and protein concentrations of the other samples were adjusted accordingly.

Enyzme assays were performed overnight at ambient temperature with native I3M from Dyer's woad (*Isatis tinctorea*), isolated according to a protocol from Pfalz et al. (2009). For each assay, 50 μL microsomes were mixed with I3M in 100 mM KPi, pH 7.4, 2 mM NADPH, 2 mM glucose-6-phosphate, and 0.1 units of glucose-6-phosphate dehydrogenase in a total volume of 200 μL. After incubation, the mixture was loaded onto DEAE-Sephadex A-25 anion-exchange columns and washed repeatedly (3× 1 mL 67% [v/v] aqueous methanol, 2× 1 mL deionized water, and 3× 1 mL MES buffer, pH 5.2). Thirty units of *H. pomatia* sulfatase (Sigma-Aldrich), suspended in 50 μL, was added to each of the columns to convert glucosinolates into their desulfo forms within 6 h of incubation at ambient temperature. Desulfo-glucosinolates were eluted from the columns with 6× 0.5 mL water, concentrated on a Genevac HT-4X Series II vacuum evaporator, and resuspended in 220 μL water. HPLC was performed with a 95-μL sample on an Agilent HP 1100 Series system with the following program: start with 1.5% solvent B (acetonitrile) and 98.5% A (water), 1 min 1.5% B, 6 min 5% B, 8 min 7% B, 18 min 21% B and 23 min 29% B, 30 min 43% B, 30.5 min 100% B, 33 min 100% B, 33.1 min 1.5% B, and 38 min 1.5% B. Finally, reaction product identity was confirmed by LC-MS.

Expression of Indole Glucosinolate Biosynthesis Genes in *N. benthamiana*

The pathway introduced consisted of the following genes (as shown in Figure 6): *CYP79B2* (Hull et al., 2000; Mikkelsen et al., 2000) and

CYP83B1 (Bak et al., 2001; Hansen et al., 2001; Naur et al., 2003) encode cytochrome P450 monooxygenases; *GSTF9*, a glutathione transferase (Wentzell et al., 2007); *GGP1*, a γ-glutamyl peptidase (Geu-Flores et al., 2009); *SUR1*, a *C-S* lyase (Mikkelsen et al., 2004); *UGT74B1*, an *S*-glucosyltransferase (Grubb et al., 2004; Gachon et al., 2005); and At *ST5a*, a sulfotransferase (Piotrowski et al., 2004).

CYP79B2 and *CYP83B1* coding sequences were amplified by PCR from *Arabidopsis* cDNA to incorporate USER-cloning sequences with primers CYP79B2f+r and with CYP81B1f+r. Each of the resulting fragments were mixed in approximately equimolar ratios with the pCAMBIA230035SU-vector, USER-treated as described previously (Nour-Eldin et al., 2006), and transformed into *E. coli*. Resulting positive clones were sequenced and transformed into *Agrobacterium tumefaciens* strain C58 by electroporation. Additional gene expression constructs harboring *GSTF9*, *GGP1*, *SUR1*, At *ST5a*, and *UGT74B1* were kindly provided by F. Geu-Flores and M.T. Nielsen. Generation of these constructs was described previously (Geu-Flores et al., 2009). *CYP81F* and *O-MT* genes were cloned into the donor vector pDONR207 using the Gateway system (Invitrogen) and then transferred into the expression vector pEARLEY201 (Earley et al., 2006).

A. tumefaciens cultures were grown at 28°C in YEP media with 50 mg/L kanamycin and 34 mg/L rifampicin. Cells were harvested by centrifugation, 10 min at 3000*g*, and resuspended to a final OD_{600} of 0.75 in 10 mM MES buffer with 10 mM $MgCl_2$ and 100 μM acetosyringone. Following a 150-min shaking incubation, 50 rpm at room temperature, the *A. tumefaciens* strains were mixed in roughly equimolar amounts and a volume of the P19 suppressor strain corresponding to approximately one-quarter of the total volume (Voinnet et al., 2003) was added. The *A. tumefaciens* suspension was injected into the leaves of 3- to 4-week-old *N. benthamiana* plants using a 1-mL syringe. Glucosinolates were extracted by applying the total extract on a Sephadex DEAE A25 column, washed with 85% methanol and water, desulfonated with a *H. pomatia* sulfatase, and eluted with water. The resulting desulfo-glucosinolates were analyzed by LC-MS as described previously (Mikkelsen et al., 2010).

In silico Analyses of Gene Expression

Publicly available microarray data were used to investigate correlations in gene expression patterns among *CYP81F* and *O-MT* genes. Data sets included 1436 Affymetrix *Arabidopsis* ATH1 microarrays (www.Arabidopsis.org), and microarray data subsets focusing on plant development, treatments with hormones, abiotic stress, or pathogens (www.weigelworld.org/resources/microarray/AtGenExpress; Schmid et al., 2005; Kilian et al., 2007). The *Arabidopsis* eFP-Browser (http://bar.utoronto.ca) was used to visualize the expression of *CYP81F* and *O-MT* genes in different tissues during plant development (Winter et al., 2007).

Statistical Programs

qRT-PCR data were analyzed with qBASE PLUS (Hellemans et al., 2007). Analysis of variance was performed with Systat Version 9 (SPSS).

Accession Numbers

Sequence data from this article can be found in the GenBank/EMBL data libraries or the Arabidopsis Genome Initiative database under accession numbers NM_119903 (*CYP81F3*; At4g37400), NM_119904 (*CYP81F4*; At4g37410), NM_119906 (*CYP81F1*; At4g37430), NM_125104 (*CYP81F2*; At5g57220), NM_101964 (*IGMT1*; At1g21100), NM_101966 (*IGMT2*; At1g21120), NM_120158 (*CYP79B2*; At4g39950), NM_119299 (*CYP83B1*; At4g31500), NM_128638 (*GSTF9*; At2g30860), NM_119199 (*GGP1*; At4g30530), NM_127622 (*SUR1*; At2g20610), NM_106070 (*AtST5a*; At1g74100), and NM_102256 (*UGT74B1*; At1g24100) . T-DNA insertions used were as follows (see Figure 1 for details): SALK_031939 (*cyp81f1*), SALK_123882 (*cyp81f2*), FLAG_140B06 (*cyp81f3/f34*), and SALK_024438 (*cyp81f4*).

Supplemental Data

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

Supplemental Figure 1. Correlations in Gene Expression: Part I.

Supplemental Figure 2. Correlations in Gene Expression: Part II.

Supplemental Figure 3. In Silico Analysis of *CYP81F* and *O-MT* Gene Expression Patterns.

Supplemental Figure 4. Engineering of *CYP81F* Genes in *N. benthamiana*: MS Data.

Supplemental Figure 5. Engineering of *CYP81F* and *O-MT* Genes in *N. benthamiana*: MS Data I.

Supplemental Figure 6. Engineering of *CYP81F* and *O-MT* Genes in *N. benthamiana*: MS Data II.

Supplemental Figure 7. Coexpression of *CYP81F* and *IGMT* Genes in *N. benthamiana*.

Supplemental Figure 8. DNA Gel Blots of T-DNA Insertion Lines.

Supplemental Table 1. Primer Sequences Used in This Work.

ACKNOWLEDGMENTS

We thank Michael Reichelt, John D'Auria, and Heiko Vogel from the Max Planck Institute for Chemical Ecology for their help with experiments. This work was partially supported by the Deutsche Forschungsgemeinschaft and by the Agence Nationale de la Recherche.

Received November 28, 2010; revised January 4, 2011; accepted January 23, 2011; published February 11, 2011.

REFERENCES

Ahn, Y.O., Shimizu, B.I., Sakata, K., Gantulga, D., Zhou, C., Bevan, D.R., and Esen, A. (2010). Scopolin-hydrolyzing β-glucosidases in roots of Arabidopsis. Plant Cell Physiol. **51:** 132–143. Erratum. Plant Cell Physiol. **51:** 339.

Albinsky, D., Sawada, Y., Kuwahara, A., Nagano, M., Hirai, A., Saito, K., and Hirai, M.Y. (2010). Widely targeted metabolomics and coexpression analysis as tools to identify genes involved in the side-chain elongation steps of aliphatic glucosinolate biosynthesis. Amino Acids **39:** 1067–1075.

Alonso, J.M., et al. (2003). Genome-wide insertional mutagenesis of *Arabidopsis thaliana.* Science **301:** 653–657.

Bak, S., Tax, F.E., Feldmann, K.A., Galbraith, D.W., and Feyereisen, R. (2001). CYP83B1, a cytochrome P450 at the metabolic branch point in auxin and indole glucosinolate biosynthesis in *Arabidopsis*. Plant Cell **13:** 101–111.

Bednarek, P., Pislewska-Bednarek, M., Svatoš, A., Schneider, B., Doubsky, J., Mansurova, M., Humphry, M., Consonni, C., Panstruga, R., Sanchez-Vallet, A., Molina, A., and Schulze-Lefert, P. (2009). A glucosinolate metabolism pathway in living plant cells mediates broad-spectrum antifungal defense. Science **323:** 101–106.

Benderoth, M., Pfalz, M., and Kroymann, J. (2009). Methylthioalkylmalate synthases: Genetics, ecology and evolution. Phytochem. Rev. **8:** 255–268.

Benderoth, M., Textor, S., Windsor, A.J., Mitchell-Olds, T., Gershenzon,

J., and Kroymann, J. (2006). Positive selection driving diversification in plant secondary metabolism. Proc. Natl. Acad. Sci. USA **103:** 9118–9123.

Brown, P.D., Tokuhisa, J.G., Reichelt, M., and Gershenzon, J. (2003). Variation of glucosinolate accumulation among different organs and developmental stages of *Arabidopsis thaliana.* Phytochemistry **62:** 471–481.

Brunaud, V., et al. (2002). T-DNA integration into the Arabidopsis genome depends on sequences of pre-insertion sites. EMBO Rep. **3:** 1152–1157.

Clarke, D.B. (2010). Glucosinolates, structures and analysis in food. Anal. Methods **2:** 310–325.

Clay, N.K., Adio, A.M., Denoux, C., Jander, G., and Ausubel, F.M. (2009). Glucosinolate metabolites required for an Arabidopsis innate immune response. Science **323:** 95–101.

D'Auria, J.C. (2006). Acyltransferases in plants: A good time to be BAHD. Curr. Opin. Plant Biol. **9:** 331–340.

Daxenbichler, M.E., Spencer, G.F., Carlson, D.G., Rose, G.B., Brinker, A.M., and Powell, R.G. (1991). Glucosinolate composition of seeds from 297 species of wild plants. Phytochemistry **30:** 2623–2638.

de Vos, M., Kriksunov, K.L., and Jander, G. (2008). Indole-3-acetonitrile production from indole glucosinolates deters oviposition by *Pieris rapae.* Plant Physiol. **146:** 916–926.

Earley, K.W., Haag, J.R., Pontes, O., Opper, K., Juehne, T., Song, K., and Pikaard, C.S. (2006). Gateway-compatible vectors for plant functional genomics and proteomics. Plant J. **45:** 616–629.

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

Fischer, H.M., Wheat, C.W., Heckel, D.G., and Vogel, H. (2008). Evolutionary origins of a novel host plant detoxification gene in butterflies. Mol. Biol. Evol. **25:** 809–820.

Freitak, D., Wheat, C.W., Heckel, D.G., and Vogel, H. (2007). Immune system responses and fitness costs associated with consumption of bacteria in larvae of *Trichoplusia ni.* BMC Biol. **5:** 56.

Gachon, C.M.M., Langlois-Meurinne, M., Henry, Y., and Saindrenan, P. (2005). Transcriptional co-regulation of secondary metabolism enzymes in *Arabidopsis*: Functional and evolutionary implications. Plant Mol. Biol. **58:** 229–245.

Geu-Flores, F., Nielsen, M.T., Nafisi, M., Møldrup, M.E., Olsen, C.E., Motawia, M.S., and Halkier, B.A. (2009). Glucosinolate engineering identifies a γ-glutamyl peptidase. Nat. Chem. Biol. **5:** 575–577.

Gigolashvili, T., Berger, B., Mock, H.P., Müller, C., Weisshaar, B., and Flügge, U.I. (2007). The transcription factor HIG1/MYB51 regulates indolic glucosinolate biosynthesis in *Arabidopsis thaliana.* Plant J. **50:** 886–901.

Glawischnig, E., Hansen, B.G., Olsen, C.E., and Halkier, B.A. (2004). Camalexin is synthesized from indole-3-acetaldoxime, a key branching point between primary and secondary metabolism in *Arabidopsis.* Proc. Natl. Acad. Sci. USA **101:** 8245–8250.

Grubb, C.D., Zipp, B.J., Ludwig-Müller, J., Masuno, M.N., Molinski, T.F., and Abel, S. (2004). Arabidopsis glucosyltransferase UGT74B1 functions in glucosinolate biosynthesis and auxin homeostasis. Plant J. **40:** 893–908.

Halkier, B.A., and Gershenzon, J. (2006). Biology and biochemistry of glucosinolates. Annu. Rev. Plant Biol. **57:** 303–333.

Hansen, C.H., Du, L.C., Naur, P., Olsen, C.E., Axelsen, K.B., Hick, A. J., Pickett, J.A., and Halkier, B.A. (2001). CYP83b1 is the oxime-metabolizing enzyme in the glucosinolate pathway in *Arabidopsis.* J. Biol. Chem. **276:** 24790–24796.

Hellemans, J., Mortier, G., De Paepe, A., Speleman, F., and Vandesompele, J. (2007). qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. Genome Biol. **8:** R19.

Hippauf, F., Michalsky, E., Huang, R., Preissner, R., Barkman, T.J., and Piechulla, B. (2010). Enzymatic, expression and structural divergences among carboxyl *O*-methyltransferases after gene duplication and speciation in *Nicotiana.* Plant Mol. Biol. **72:** 311–330.

Hirai, M.Y., et al. (2007). Omics-based identification of Arabidopsis Myb transcription factors regulating aliphatic glucosinolate biosynthesis. Proc. Natl. Acad. Sci. USA **104:** 6478–6483.

Hull, A.K., Vij, R., and 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.

Kilian, J., Whitehead, D., Horak, J., Wanke, D., Weinl, S., Batistic, O., D'Angelo, C., Bornberg-Bauer, E., Kudla, J., and Harter, K. (2007). The AtGenExpress global stress expression data set: Protocols, evaluation and model data analysis of UV-B light, drought and cold stress responses. Plant J. **50:** 347–363.

Kim, J.H., and Jander, G. (2007). *Myzus persicae* (green peach aphid) feeding on Arabidopsis induces the formation of a deterrent indole glucosinolate. Plant J. **49:** 1008–1019.

Kim, J.H., Lee, B.W., Schroeder, F.C., and Jander, G. (2008). Identification of indole glucosinolate breakdown products with antifeedant effects on *Myzus persicae* (green peach aphid). Plant J. **54:** 1015–1026.

Kliebenstein, D.J. (2009). A quantitative genetics and ecological model system: understanding the aliphatic glucosinolate biosynthetic network via QTLs. Phytochem. Rev. **8:** 243–254.

Kliebenstein, D.J., Kroymann, J., Brown, P., Figuth, A., Pedersen, D., Gershenzon, J., and Mitchell-Olds, T. (2001). Genetic control of natural variation in Arabidopsis glucosinolate accumulation. Plant Physiol. **126:** 811–825.

Kliebenstein, D.J., Kroymann, J., and Mitchell-Olds, T. (2005). The glucosinolate-myrosinase system in an ecological and evolutionary context. Curr. Opin. Plant Biol. **8:** 264–271.

Kozak, M. (1987). An analysis of 5′-noncoding sequences from 699 vertebrate messenger RNAs. Nucleic Acids Res. **15:** 8125–8148.

Kozak, M. (1990). Downstream secondary structure facilitates recognition of initiator codons by eukaryotic ribosomes. Proc. Natl. Acad. Sci. USA **87:** 8301–8305.

Kozak, M. (1991). An analysis of vertebrate mRNA sequences: Intimations of translational control. J. Cell Biol. **115:** 887–903.

Kroymann, J., Donnerhacke, S., Schnabelrauch, D., and Mitchell-Olds, T. (2003). Evolutionary dynamics of an *Arabidopsis* insect resistance quantitative trait locus. Proc. Natl. Acad. Sci. USA **100** (suppl. 2): 14587–14592.

Kroymann, J., Textor, S., Tokuhisa, J.G., Falk, K.L., Bartram, S., Gershenzon, J., and 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.

Lam, K.C., Ibrahim, R.K., Behdad, B., and Dayanandan, S. (2007). Structure, function, and evolution of plant *O*-methyltransferases. Genome **50:** 1001–1013.

Mikkelsen, M.D., Hansen, C.H., Wittstock, U., and 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., Naur, P., and 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.

Mikkelsen, M.D., Olsen, C.E., and Halkier, B.A. (2010). Production of the cancer-preventive glucoraphanin in tobacco. Mol. Plant **3:** 751–759.

Naur, P., Petersen, B.L., Mikkelsen, M.D., Bak, S., Rasmussen, H., Olsen, C.E., and 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.

Nitz, I., Berkefeld, H., Puzio, P.S., and Grundler, F.M.W. (2001). *Pyk10*, a seedling and root specific gene and promoter from *Arabidopsis thaliana.* Plant Sci. **161:** 337–346.

Noel, J.P., Dixon, R.A., Pichersky, E., Zubieta, C., and Ferrer, J.L. (2003). Structural, functional and evolutionary basis for methylation of plant small molecules. In Recent Advances in Phytochemistry, Vol. 37. Integrative Phytochemistry: from Ethnobotany to Molecular Ecology, J.T. Romero, ed (Oxford, U.K.: Elsevier Science Ltd), pp. 37–58.

Nour-Eldin, H.H., Hansen, B.G., Nørholm, M.H.H., Jensen, J.K., and Halkier, B.A. (2006). Advancing uracil-excision based cloning towards an ideal technique for cloning PCR fragments. Nucleic Acids Res. **34:** e122.

Obayashi, T., Hayashi, S., Saeki, M., Ohta, H., and Kinoshita, K. (2009). ATTED-II provides coexpressed gene networks for Arabidopsis. Nucleic Acids Res. **37** (Database issue): D987–D991.

Pfalz, M., Vogel, H., and Kroymann, J. (2009). The gene controlling the *indole glucosinolate modifier1* quantitative trait locus alters indole glucosinolate structures and aphid resistance in *Arabidopsis.* Plant Cell **21:** 985–999.

Pfalz, M., Vogel, H., Mitchell-Olds, T., and Kroymann, J. (2007). Mapping of QTL for resistance against the crucifer specialist herbivore *Pieris brassicae* in a new Arabidopsis inbred line population, Da(1)-12 x Ei-2. PLoS ONE **2:** e578.

Pinçon, G., Maury, S., Hoffmann, L., Geoffroy, P., Lapierre, C., Pollet, B., and Legrand, M. (2001). Repression of *O*-methyltransferase genes in transgenic tobacco affects lignin synthesis and plant growth. Phytochemistry **57:** 1167–1176.

Piotrowski, M., Schemenewitz, A., Lopukhina, A., Müller, A., Janowitz, T., Weiler, E.W., and 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.

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

Sawada, Y., Kuwahara, A., Nagano, M., Narisawa, T., Sakata, A., Saito, K., and Hirai, M.Y. (2009). Omics-based approaches to methionine side chain elongation in Arabidopsis: Characterization of the genes encoding methylthioalkylmalate isomerase and methylthioalkylmalate dehydrogenase. Plant Cell Physiol. **50:** 1181–1190.

Schmid, M., Davison, T.S., Henz, S.R., Pape, U.J., Demar, M., Vingron, M., Schölkopf, B., Weigel, D., and Lohmann, J.U. (2005). A gene expression map of *Arabidopsis thaliana* development. Nat. Genet. **37:** 501–506.

Sherameti, I., Venus, Y., Drzewiecki, C., Tripathi, S., Dan, V.M., Nitz, I., Varma, A., Grundler, F.M.W., and Oelmüller, R. (2008). PYK10, a β-glucosidase located in the endoplasmatic reticulum, is crucial for the beneficial interaction between *Arabidopsis thaliana* and the endophytic fungus *Piriformospora indica.* Plant J. **54:** 428–439.

Sønderby, I.E., Geu-Flores, F., and Halkier, B.A. (2010). Biosynthesis of glucosinolates—Gene discovery and beyond. Trends Plant Sci. **15:** 283–290.

Sun, J.Y., Sønderby, I.E., Halkier, B.A., Jander, G., and de Vos, M. (2009). Non-volatile intact indole glucosinolates are host recognition cues for ovipositing *Plutella xylostella.* J. Chem. Ecol. **35:** 1427–1436.

Voinnet, O., Rivas, S., Mestre, P., and Baulcombe, D. (2003). An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. Plant J. **33:** 949–956.

Wentzell, A.M., Rowe, H.C., Hansen, B.G., Ticconi, C., Halkier, B.A., and Kliebenstein, D.J. (2007). Linking metabolic QTLs with network and *cis*-eQTLs controlling biosynthetic pathways. PLoS Genet. **3:** 1687–1701.

Wink, M. (2010). Biochemistry, physiology and ecological functions of secondary metabolites. In Annual Plant Reviews, Vol. 40. Biochemistry of Plant Secondary Metabolism, 2nd ed, M. Wink, ed (Hoboken, NJ: Wiley-Blackwell), pp. 1–19.

Winter, D., Vinegar, B., Nahal, H., Ammar, R., Wilson, G.V., and Provart, N.J. (2007). An "Electronic Fluorescent Pictograph" browser for exploring and analyzing large-scale biological data sets. PLoS ONE **2:** e718.

Wu, S., Watanabe, N., Mita, S., Dohra, H., Ueda, Y., Shibuya, M., and Ebizuka, Y. (2004). The key role of phloroglucinol O-methyltransferase in the biosynthesis of Rosa chinensis volatile 1,3,5-trimethoxybenzene. Plant Physiol. **135:** 95–102.