Modularity of Plant Metabolic Gene Clusters: A Trio of Linked Genes That Are Collectively Required for Acylation of Triterpenes in Oat

Sam T. Mugford, Thomas Louveau, Rachel Melton, Xiaoquan Qi, Saleha Bakht, Lionel Hill, Tetsu Tsurushima, Suvi Honkanen, Susan J. Rosser, George P. Lomonossoff, and Anne Osbourn

INTRODUCTION

Plants synthesize a diverse range of specialized metabolites (also commonly referred to as natural products), many of which have important roles in protection against disease, herbivory, and/or abiotic stress (Dixon, 2001; Osbourn and Lanzotti, 2009). Oats (Avena spp) produce antimicrobial triterpene glycosides (avenacin) in their roots that protect against soil-borne pathogens (Hostettmann and Marston, 1995; Papadopoulo et al., 1999; Osbourn et al., 2011). Although a diverse array of structurally varied triterpene glycosides is found in dicots, the ability to synthesize glycosylated triterpenes is rare in the cereals and grasses. Avenacins exhibit some unusual structural features that are not shared by other plant triterpene glycosides. Most notably, they are acylated at the carbon-21 (C-21) position with either N-methyl anthranilate or benzoate (Figure 1). A recent review of structure-function relationships in triterpenoids indicates that C-21 acylation is important in determining the biological activity of some triterpenoids (Podolak et al., 2010). Understanding the mechanisms underlyling this type of modification has potential to provide powerful biotechnological tools for the modification of triterpene structure and function in different biological systems.

Operon-like gene clusters are an emerging phenomenon in the field of plant natural products. The genes encoding some of the best-characterized plant secondary metabolite biosynthetic pathways are scattered across plant genomes. However, an increasing number of gene clusters encoding the synthesis of diverse natural products have recently been reported in plant genomes. These clusters have arisen through the neo-functionalization and relocation of existing genes within the genome, and not by horizontal gene transfer from microbes. The reasons for clustering are not yet clear, although this form of gene organization is likely to facilitate co-inheritance and co-regulation. Oats (Avena spp) synthesize antimicrobial triterpenoids (avenacins) that provide protection against disease. The synthesis of these compounds is encoded by a gene cluster. Here we show that a module of three adjacent genes within the wider biosynthetic gene cluster is required for avenacin acylation. Through the characterization of these genes and their encoded proteins we present a model of the subcellular organization of triterpenoid biosynthesis.

Acylation orchestrates the amalgamation of complex chemical groups with different biochemical origins into single compounds and so offers great versatility in modulating the form and function of plant natural products (D’Auria, 2006). Plant acyltransferase enzymes have been identified that catalyze the transfer of a wide range of acyl groups onto a diverse set of acceptor substrates belonging to different classes of natural product. These enzymes belong to two unrelated families: the BAHD acyltransferases, which use CoA-thioesters as the acyl donor (Yang et al., 1997; Dudareva et al., 1998; Fujiwara et al., 1998; St-Pierre et al., 1998; Walker et al., 2002; Luo et al., 2007, 2009; Grienenberger et al., 2009); and the serine carboxypeptidase-like (SCPL) acyltransferases, which use O-Glc esters as acyl donors. The SCPL-acyltransferases belong to the serine carboxypeptidase family but have acquired different functions to these ancestral proteases (Lehfeldt et al., 2000; Li and Steffens, 2000; Shirley et al., 2001; Fraser et al., 2007; Weier et al., 2008; Shirley and Chapelle, 2003; Milkowski and Strack, 2004; Milkowski et al., 2004; Baumert et al., 2005; Stehle et al., 2008; Mugford et al., 2009; Mugford and Osbourn, 2010; Mugford and Milkowski, 2012). Based on the available evidence, it seems likely that the two different families of acyltransferases are also distinguished by their subcellular localizations. The Arabidopsis thaliana sinapoyl-malate SCPL-acyltransferase SMT has been localized to the vacuole (Hause et al., 2002), and other members of this family are likewise predicted to be targeted to the vacuole in Arabidopsis and other plants (Fraser et al., 2005). Conversely, members of the BAHD family have either been shown to be or are predicted to be cytosolic (Fujiwara et al., 1998; D’Auria, 2006; Yu et al., 2008).

We previously established a collection of avenacin-deficient mutants of the diploid oat species Avena strigosa. The major
Avenacin A-1, which is acylated with N-methyl anthranilate (Crombie et al., 1984).

Previously, we identified and sequenced a BAC contig spanning four genes within the avenacin gene cluster: Sad1, Sad2, Sad7, and Sad10 (Qi et al., 2006; Mugford et al., 2009; Owatworakit et al., 2012). Further extension of this BAC contig identified a gene predicted to encode an S-adenosyl-methionine–dependent methyltransferase (MT1), which lies 45 kb from the UGT74H5 (Sad10) gene (Figure 2A). Avenacins A-1 and B-1 are both acylated with N-methyl anthranilate, implicating anthranilate methylation in the avenacin biosynthetic pathway. Phylogenetic analysis shows that MT1 lies within a clade of methyltransferases that consists mainly of O-methyltransferase enzymes with roles in secondary metabolism (Figure 2B; see Supplemental Data Set 1 online), although two characterized members of this clade (Limonium latifolium β-alanine-N-methyltransferase and Ruta graveolens anthranilate-N-methyltransferase) are known to have N-methyltransferase activity (Lisacombe and Facchini, 2007; Rohde et al., 2008), consistent with a possible role for MT1 as an N-methyltransferase.

Avenacins are synthesized in the epidermal cells of the root tip, and the expression of genes that encode the previously characterized steps in the biosynthetic pathway is restricted to these cells (Qi et al., 2006; Wegel et al., 2009; Mugford et al., 2009). Quantitative PCR analysis revealed that the MT1 transcript is most abundant in whole root and root tip samples, low in the upper root, and not detectable in young leaves (Figure 2C). Immunoblot analysis detected MT1 in extracts from the roots but not the leaves of wild-type A. strigosa seedlings, as is the case for other avenacin biosynthetic enzymes (Figure 3, lanes 1 and 2). Furthermore, immunostaining of root tip tissue sections with specific antisera raised against As-MT1 indicates that the protein is restricted to the epidermal cells of the root tip (Figure 2D). A similar pattern of staining is also observed using antisera raised against the As-SCPL1 protein. Signals were not observed for sections stained with the corresponding preimmune sera (see Supplemental Figure 1 online). Collectively, these results are consistent with a role for As-MT1 in avenacin biosynthesis.

**RESULTS**

**Three Genes Encoding a Methyltransferase, Glucosyltransferase, and an Acyltransferase Are Adjacent within the Avenacin Gene Cluster**

Previously, we identified and sequenced a BAC contig spanning four genes within the avenacin gene cluster: Sad1, Sad2, Sad7, and Sad10 (Qi et al., 2006; Mugford et al., 2009; Owatworakit et al., 2012). Further extension of this BAC contig identified a gene predicted to encode an S-adenosyl-methionine–dependent methyltransferase (MT1), which lies 45 kb from the UGT74H5 (Sad10) gene (Figure 2A). Avenacins A-1 and B-1 are both acylated with N-methyl anthranilate, implicating anthranilate methylation in the avenacin biosynthetic pathway. Phylogenetic analysis shows that MT1 lies within a clade of methyltransferases that consists mainly of O-methyltransferase enzymes with roles in secondary metabolism (Figure 2B; see Supplemental Data Set 1 online), although two characterized members of this clade (Limonium latifolium β-alanine-N-methyltransferase and Ruta graveolens anthranilate-N-methyltransferase) are known to have N-methyltransferase activity (Lisacombe and Facchini, 2007; Rohde et al., 2008), consistent with a possible role for MT1 as an N-methyltransferase. Avenacins are synthesized in the epidermal cells of the root tip, and the expression of genes that encode the previously characterized steps in the biosynthetic pathway is restricted to these cells (Qi et al., 2006; Wegel et al., 2009; Mugford et al., 2009). Quantitative PCR analysis revealed that the MT1 transcript is most abundant in whole root and root tip samples, low in the upper root, and not detectable in young leaves (Figure 2C). Immunoblot analysis detected MT1 in extracts from the roots but not the leaves of wild-type A. strigosa seedlings, as is the case for other avenacin biosynthetic enzymes (Figure 3, lanes 1 and 2). Furthermore, immunostaining of root tip tissue sections with specific antisera raised against As-MT1 indicates that the protein is restricted to the epidermal cells of the root tip (Figure 2D). A similar pattern of staining is also observed using antisera raised against the As-SCPL1 protein. Signals were not observed for sections stained with the corresponding preimmune sera (see Supplemental Figure 1 online). Collectively, these results are consistent with a role for As-MT1 in avenacin biosynthesis.

MT1 Is Required for N-Methylation of Avenacin A-1 and Correlates to Sad9, a Locus Previously Defined by Mutation

We previously isolated a collection of avenacin-deficient mutants of A. strigosa that were identified from a screen for loss of UV fluorescence in the root tip (Papadopoulou et al., 1999; Qi et al.,
Figure 2. A Gene Encoding a Predicted Methyltransferase Located within the Avenacin Biosynthetic Gene Cluster Is Implicated in Avenacin Synthesis.

(A) BAC contig showing the avenacin gene cluster and the location of the Sad9 gene, which is predicted to encode a methyltransferase. Gene names are indicated above and encoded proteins below.

(B) Neighbor-joining phylogeny of functionally characterized plant natural product S-adenosyl-Met–dependent methyltransferase enzymes. N-methyltransferase enzymes are indicated with asterisks. The phylogeny includes the sequences used by Liscombe and Facchini, 2007; the MT1 protein (indicated by an arrow); the most similar sequences to As-MT1 encoded by the genomes of Arabidopsis (At4g35160), rice (Oryza sativa; Os9g17560), Brachypodium distachyon (Bd2g19830), and sorghum (Sorghum bicolor; Sb07g024270); and the R. graveolens anthranilate methyltransferase (Rg...
LC-MS analysis revealed that DMA is found not only in
leaves of N-
aceylantranilate (blue) compared with that of anthranilate
methyl anthranilate (orange). Immunoblot analysis of soluble protein extracts from 3-d-old A. strigosa
seedlings, probed with antiserum raised against MT1, UGT74H5, or SCPL1. The bottom panel shows Coomassie blue staining of a replicate gel.

2006). Liquid chromatography–mass spectrometry (LC-MS) analysis of this mutant collection identified four independent mutants that lacked avenacin A-1 but accumulated a related compound that is acylated with anthranilate rather than N-methyl anthranilate (des-methyl avenacin A-1 [DMA]) (Figure 4A; see Supplemental Figure 2 online). These four mutant lines (#1475, #1310, #961, and #841) were shown to be allelic. The roots of early avenacin pathway mutants (sad1 and sad2) (Papadopoulou et al., 1999; Qi et al., 2004, 2006; Qin et al., 2010) have little or no fluorescence, while those of sad7 have weak blue autofluorescence due to the accumulation of the acyl donor N-methyl anthraniloyl-O-Glc (Mugford et al., 2009). By contrast, roots of sad9 mutants have a weak purple fluorescence under UV illumination. The fluorescence of extracts from wild-type roots peaks at 440 nm compared with 410 nm from the sad9 mutant. This is consistent with the fluorescence of N-methyl anthranilate (blue) compared with that of anthranilate (purple). LC-MS analysis revealed that DMA is found not only in sad9 mutants, but also in wild-type A. strigosa roots (Figure 4). sad9 mutants lack N-methyl anthraniloyl-O-Glc but accumulate a fluorescent compound with mass spectrum and fragmentation pattern consistent with anthraniloyl-O-glucose (Figure 4A; see Supplemental Figure 2 online). Quantification of acylated avenacins in a selection of mutants showed that DMA accumulates in the sad9 mutants at lower levels (109.0 ± 29.9 pmol/mg fresh weight in #961; 189.2 ± 26.5 pmol/mg fresh weight in #1310) than is found for avenacin A-1 in the wild type (365.3 ± 8.9 pmol/mg fresh weight) (values are means of four replicate samples ± se). In addition, the sad9 mutants accumulate des-acyl avenacin, which is also found in the sad7 mutants (Figure 4), and the nonfluorescent avenacins A-2 and B-2, which are acylated with benzoate and are also present at similar levels in the wild type (see Supplemental Table 1 online).

The failure of sad9 mutants to methylate avenacin and the presence of MT1 (encoding a predicted methyltransferase) in the avenacin gene cluster suggests that Sad9 and MT1 may be synonymous. DNA sequence analysis of MT1 in the sad9 mutants revealed nonsynonymous point mutations in all four alleles (#841, C998T and A333V; #961, C89T and S30F; #1310, G978A and W326Stop; #1475, C235T and R79W) that alter single amino acids or in the case of #1310 introduce a stop codon, indicating that Sad9 is likely to encode MT1. These data suggest a pathway for acylation of avenacins as shown in Figure 4B that involves N-methylation of anthranilate by MT1 (SAD9), glucosylation of N-methyl anthranilate by UGT74H5 (SAD10), and the transfer of the acyl group to the triterpene backbone by SCPL1 (SAD7). In the absence of MT1, SCPL1 is apparently still able to acylate avenacins using anthraniloyl-O-Glc as the acyl donor instead of N-methyl anthraniloyl-O-Glc. Previously we have shown that sad1, sad2, and sad7 mutants of A. strigosa have increased susceptibility to soil-borne fungal pathogens (Papadopoulou et al., 1999; Qi et al., 2004, 2006; Mugford et al., 2009). Pathogenicity tests with Gaeumannomyces graminis var tritici, the causal agent of take-all disease of cereals, revealed that, like sad1, sad2, and sad7 mutants, sad9 mutants are more susceptible to infection than wild-type seedlings (Figure 5). The increased susceptibility of sad7 and sad9 mutants to disease may be because correct acylation of avenacins is required for antifungal activity or, alternatively, because the unacylated/incorrectly acylated avenacins accumulate at lower levels in the mutants. We measured the total concentrations of avenacins and des-acyl avenacins in wild-type and sad7 mutant oat roots and found that the levels of unacylated avenacins in the sad7 mutant roots were comparable to the levels of acylated avenacins in the wild type (see Supplemental Figure 3 online), indicating that the acyl group is likely to be important in protecting against disease. Furthermore, bioassays

Figure 3. Detection of MT1, UGT74H5, and SCPL1 in Oat Roots. Immunoblot analysis of soluble protein extracts from 3-d-old A. strigosa seedlings, probed with antiserum raised against MT1, UGT74H5, or SCPL1. The bottom panel shows Coomassie blue staining of a replicate gel. Protein extracts were prepared from leaves and roots of 3-d-old A. strigosa wild-type (WT) seedlings and from roots of previously characterized sad1 (mutant line #610) and sad7 (#616) mutants (Papadopoulou et al., 1999; Mugford et al., 2009) and sad9 mutants (#1475, 1310, 961, and 841) (this article). Bands corresponding to the predicted size of the MT1 and UGT74H5 proteins were detected. SCPL1 is posttranslationally cleaved into a large and a small subunit corresponding to the bands detected at 29 and 19 kD and a processing intermediate (33 kD) (Mugford et al., 2009).

Figure 2. (continued).
ANMT; Rohde et al., 2008). Details of these other sequences are provided in Methods, and the full protein sequence alignment can be found in Supplemental Data Set 1 online. Bootstrap values (percentage of 1000 replicates) are shown for key branches. Bar indicates 0.05 substitutions per site.
(C) Sad9 (AsmM77) transcript abundance in young leaves, whole roots, root tips (0.5 mm), and roots minus tips, measured by quantitative PCR (transcript abundance expressed relative to elongation factor 1-α); values are means (n = 3 ± se).
(D) Immunostaining of root tip sections from 3-d-old wild-type A. strigosa seedlings. Sections were stained with (from left to right) anti-AsMT1 antisera, anti-AsSCPL1 antisera, or 4',6-diamidino-2-phenylindole (DAPI). A bright-field image and an overlay of the MT1, SCPL1, and DAPI images are also shown. Controls probed with preimmune sera are shown in Supplemental Figure 1 online.
with extracts from wild-type and sad7 mutant root tips separated by thin layer chromatography (TLC) show that while avenacin A-1 in the wild-type sample strongly inhibits the growth of the indicator fungus Colletotrichum orbiculare, the presence of similar amounts of desacyl avenacin A in the sad7 mutant extract has no such inhibitory effect (see Supplemental Figure 4 online). Thus, the acyl group is important for biological activity. The presence in the sad9 mutants of benzoylated forms of avenacin in addition to DMA (see Supplemental Table 1 online) complicates efforts to establish the importance of the methyl}

**Figure 4.** sad9 Mutants Are Blocked in the Methylation of Avenacin Acyl-Groups.
(A) Analysis of wild-type and mutant (sad1, sad7, and sad9) root extracts by fluorescence LC-MS. Fluorescence detection shows the absence of avenacin A-1 (A1) in all mutants and the accumulation of DMA in sad9 mutants. N-methyl anthraniloyl-O-Glc (NMA-glc) is not detectable in root extracts from sad9 mutants, but anthraniloyl-O-Glc (Anth-glc) is detected at low levels. Similar metabolite profiles were observed in the other sad9 mutants analyzed (see Supplemental Figure 2 online). Confirmation of the identity of these compounds by mass spectrometry is presented in Supplemental Figure 2 online. LC-MS shows the accumulation of des-acyl avenacin A (DAA; extracted ion chromatogram for mass-to-charge ratio = 983 to 985) in the sad7 mutant, as previously shown by Mugford et al. (2009). AU, arbitrary units.
(B) Proposed pathway for acylation of avenacin A-1. N-methyl anthranilate (NMA) is synthesized by the methylation of anthranilate (Anth) by MT1. The glucosyltransferase UGT74H5 glucosylates N-methyl anthranilate to give the activated acyl donor substrate N-methyl anthraniloyl-O-Glc (NMA-Glc). UGT74H5 has a clear preference for N-methyl anthranilate but also has weak activity toward anthranilate to give anthraniloyl-O-Glc (Anth-Glc) (Owatworakit et al., 2012). In wild-type A. strigosa, the SCPL1 acyltransferase uses NMA-Glc as the acyl donor to give avenacin A-1 (A1). In sad9 mutants (lacking MT1), the triterpenoid backbone is acylated with anthranilate instead of N-methyl anthranilate to give DMA.
group for biological activity. However, a thin layer chromatography bioassay with extracts from sad9 mutant roots clearly shows zones of C. orbiculare growth inhibition associated with both of the major forms of avenacin present in this mutant, desmethyl avenacin A-1 and avenacin A-2, indicating that desmethyl avenacin A-1 does possess antifungal activity (see Supplemental Figure 4 online). It is therefore likely that the increased susceptibility of sad9 mutants to fungal infection is due to the overall lower concentrations of acylated avenacins relative to the wild type (see Supplemental Table 1 online).

MT1 was detected in protein preparations from the roots of wild-type A. strigosa seedlings and in sad1 and sad7 mutant lines. However, it was not detected in the four sad9 mutants (Figure 3). The sad9 mutants were isolated from a forward screen for loss-of-function mutations. Mutant #1310 has a premature stop codon, while, based on comparison with the structure of the Medicago truncatula isoflavone O-methyltransferase (the most similar sequence to As-MT1 for which a crystal structure is available; National Center for Biotechnology Information accession number 2QYO_A), the amino acid substitutions in the other three mutants are all likely to disrupt α-helices integral to the structure of the protein. All four sad9 mutants had wild-type levels of the SCPL1 (SAD7) and UGT74H5 (SAD10) proteins. Interestingly, the levels of SCPL1 (SAD7) protein were substantially reduced in sad1 mutants, which are blocked in the first committed step in avenacin biosynthesis (Qi et al., 2004) (shown for sad1 mutant #610 in Figure 3). Examination of the levels of SCPL1 (SAD7) protein in a wider range of early pathway mutants (Qi et al., 2004, 2006; Qin et al., 2010) revealed that the SCPL1 protein level is reduced or abolished in all sad7 and sad2 mutants, while the levels of the MT1 and UGT74H5 proteins were not substantially altered in these mutants (see Supplemental Figure 5 online). The levels of the Sad7 transcript were not significantly altered in any of the mutants (see Supplemental Figure 5 online), suggesting that SCPL1 may be posttranscriptionally regulated. A scenario consistent with these data is that the triterpene backbone is required for the accumulation of SCPL1 protein, since mutants that are unable to synthesize this backbone fail to accumulate SCPL1.

**MT1 Encodes an Anthranilate N-Methyltransferase**

MT1 was expressed in *Escherichia coli* with a C-terminal 6XHis epitope tag and purified by metal affinity chromatography. Accumulation of des-methyl avenacin and anthraniloyl-O-Glc in the sad9 mutants indicates that the likely substrate for this enzyme in planta is either anthranilate or anthraniloyl-O-Glc. We found no detectable activity of the recombinant protein against anthraniloyl-O-Glc at a range of substrate concentrations (0.5 to 5 mM), while under the same conditions, MT1 effectively catalyzed the formation of N-methyl anthranilate from anthranilate. Kinetic analysis of the MT1 enzyme in vitro showed that the enzyme has a strong affinity for anthranilic acid (K_m = 5.51 ± 0.11 μM), with a pH optimum of 8.0, and a high catalytic efficiency (k_cat/K_m = 5.55 ± 0.16 nM⁻¹ s⁻¹), which compare favorably with the reported kinetic properties of other plant methyl transferases that are active toward small phenolic compounds (Wang and Pichersky, 1999; Gang et al., 2002; Rohde et al., 2008; Jonczyk et al., 2008). Thus, As-MT1 catalyzes the first committed step in the synthesis of the N-methyl anthraniloyl acyl donor that is used by the As-SCPL1 acyltransferase, namely, methylation of the primary shikimate pathway intermediate anthranilate. This is consistent with the proposed pathway in Figure 4B and with the catalytic properties of As-UGT74H5, which has a strong affinity for N-methyl anthranilate compared with anthranilate (Owatworakit et al., 2012).

**Coexpression of As-MT1 and As-UGT74H5 in Nicotiana benthamiana Leaves Results in the Accumulation of N-Methyl Anthraniloyl-O-Glc**

The above evidence coupled with the previous characterization of As-SCPL and As-UGT74H5 (Mugford et al., 2009; Owatworakit et al., 2012) implicates As-MT1, As-UGT74H5, and As-SCPL1 in the synthesis and transfer of the N-methyl anthraniloyl acyl group onto the triterpene backbone. To further test the roles of these enzymes, we transiently coexpressed the As-MT1, As-UGT74H5, and As-SCPL1 cDNAs in N. benthamiana leaves using a cowpea mosaic virus (CPMV)-derived expression system (Sainsbury et al., 2002; Rohde et al., 2008; Jonczyk et al., 2008). Thus, As-MT1 catalyzes the first committed step in the synthesis of the N-methyl anthraniloyl acyl donor that is used by the As-SCPL1 acyltransferase, namely, methylation of the primary shikimate pathway intermediate anthranilate. This is consistent with the proposed pathway in Figure 4B and with the catalytic properties of As-UGT74H5, which has a strong affinity for N-methyl anthranilate compared with anthranilate (Owatworakit et al., 2012).

![Figure 5. sad9 Mutants Have Enhanced Disease Susceptibility.](image-url)

Seedlings were inoculated with the take-all fungus (G. graminis var tritici strain T2) and scored for disease symptoms 3 weeks later as described by Papadopoulou et al. (1999). Dark-brown/black lesions on the roots (arrowed) are typical symptoms of infection. The wild type (WT) (A), sad1 mutant #610 (B), sad7 mutant #616 (C), and sad9 mutant #1475 (D). Images are representative of six to 10 biological replicates across two independent experiments. Bars = 0.25 cm.
Figure 6. Coexpression of As-MT1 and As-UGT74H5 in Tobacco Leaves Results in the Accumulation and Sequestration of N-Methyl Anthraniloyl-O-Glc.

(A) Immunoblot analysis of soluble protein extracts of N. benthamiana leaves probed with antisera raised against MT1, UGT74H5, and SCPL1. The bottom panel shows a replicate gel stained with Coomassie blue. Lanes were loaded with protein extracted from leaves 6 d after infiltration with Agrobacterium tumefaciens strains containing (from left to right) the P19 silencing suppressor alone (P19) or in combination with the CPMV empty vector (EV); vectors containing the coding sequences of MT1, UGT74H5, or SCPL1 alone or coinfiltrated in combination as indicated.
Table 2 and Supplemental Figure 6 online) and are likely to have osm or malate residue, respectively (Figure 6C; see Supplemental Table 2 online). Expression of UGT74H5 alone did not lead to the production of detectable levels of any fluorescent compounds. Expression of the SCPL1 acyltransferase together with MT1 and UGT74H5 did not lead to accumulation of any additional detectable fluorescent compounds, indicating that SCPL1 is unlikely to be able to acylate endogenous compounds in N. benthamiana, which is perhaps not surprising since there are no reports of triterpenoid saponins or related compounds in this species. Further analysis by LC-MS failed to identify any other compounds that accumulated in an SCPL1-dependent manner. The expression of MT1 alone led to the accumulation of low levels of N-methyl anthranilate and N-methyl anthraniloyl-O-Glc, suggesting that N. benthamiana possesses enzymes capable of glucosylating N-methyl anthranilate at low levels, presumably as a side reaction of an endogenous glucosyltransferase. Also, two other fluorescent compounds were found to accumulate in leaves expressing MT1 and UGT74H5 together (and at lower levels in leaves expressing MT1 alone). These were identified by LC-MS as derivatives of N-methyl anthraniloyl-O-Glc, apparently modified by the addition of a hexose or malate residue, respectively (Figure 6C; see Supplemental Table 2 and Supplemental Figure 6 online) and are likely to have arisen as a result of endogenous N. benthamiana enzyme activities acting on N-methyl anthraniloyl-O-Glc. Confocal microscopy of N. benthamiana leaves suggests that the fluorescent material generated by coexpression of MT1 and UGT74H5 is located in the vacuole (Figure 6D).

Subcellular Localization of the Components of the Avenacin Acylation Pathway

Avenacins accumulate in the vacuoles of the epidermal cells of oat root tips (Mylonas et al., 2008). However, the primary shikimate pathway leading to the synthesis of the precursors of the acyl groups is expected to be cytosolic or, more likely, plastidial (Radwanski and Last, 1995; Schmid and Amrhein, 1995). Immunogold labeling of epidermal cells in ultrathin sections of 3-d-old oat root tips was performed using antisera raised against SCPL1, UGT74H5, or MT1 (Figure 7; see Supplemental Table 3 online). The density of gold particles associated with different cellular compartments was determined across replicate samples probed with either specific antisera or the corresponding preimmune sera (see Supplemental Table 3 online). The MT1 antisera principally labeled the cytosol in wild-type oat root epidermal cells. Signal was not detected in the control sad9 mutant line #961 or in samples probed with the corresponding preimmune sera (Figures 7A and 7B; see Supplemental Table 3 online). Sections taken from N. benthamiana leaves expressing MT1 protein were also probed with the anti-AsMT1 antisera (Figures 7C and 7D; see Supplemental Table 3 online). These were found to show strong labeling mainly in the cytosol, whereas empty vector control samples and samples probed with the preimmune sera did not give significant signals (see Supplemental Table 3 online).

Antisera raised against UGT74H5 labeled the cytoplasm and also the vacuole in wild-type oat root samples (Figure 7E) but not in samples probed with the corresponding preimmune sera (see Supplemental Table 3 online) (mutants lacking the UGT74H5 protein are not available). It should be noted that two close homologs of UGT74H5 are also expressed in oat roots and cross-react with the anti-AsUGT74H5 antisera (Owatworakit et al., 2012). To determine the localization of UGT74H5 in the absence of these other proteins, sections of N. benthamiana leaves expressing UGT74H5 were probed with the anti-AsUGT74H5 antisera. Strong labeling of the cytosol but not the vacuole was observed (Figures 7F and 7G; see Supplemental Table 3 online). This labeling was not seen in empty vector controls or in samples probed with the corresponding preimmune sera (see Supplemental Table 3 online). This suggests that UGT74H5 may also be restricted to the cytosol in oat roots and that the signal obtained from the vacuole in oat roots might be attributable to one of the close homologs of UGT74H5.

The SCPL1 antisera specifically labeled the vacuole in wild-type oat roots. No labeling was seen in sad7 mutants, which lack SCPL1, or in samples probed with preimmune serum (Figures 7H and 7I; see Supplemental Table 3 online). Taken together,
our localization experiments suggest a model for the organization of the pathway as shown in Figure 8.

**DISCUSSION**

We have shown that three consecutive genes within the avenacin gene cluster are involved in the acylation of avenacins at the C-21 position and that the failure to correctly acylate avenacins results in enhanced disease susceptibility. This is in line with the finding that C-21 acylation is a common feature of many of the most cytotoxic triterpenoid glycosides, suggesting that it is a key contributor to the biological activity of these compounds (Podolak et al., 2010). The identification of a cassette of genes for triterpene acylation is thus of practical significance and opens up opportunities for triterpene modification using metabolic engineering approaches.

The biochemical characterization and localization data presented here suggest a model for the subcellular organization of the
The shikimate pathway intermediate anthranilate is methylated and subsequently glucosylated in the cytosol, producing the activated acyl donor N-methyl anthraniloyl-O-glucose. This compound is then transported into the vacuole where it serves as a substrate for SCPL. Evidence for the transport of N-methyl anthraniloyl-O-Glc comes from sad7 mutants, which accumulate this compound, where it can be seen that the fluorescence originates from the vacuole (see Supplemental Figure 7 online). The capacity to transport the acyl donor substrate into the vacuole is also present in *N. benthamiana* (Figure 6D). The identity of the transporter(s) involved remains to be established. Previous reports of the localization of members of the S-adenosyl Met-dependent methyltransferase and family 1 glucosyltransferase families indicate that most of these enzymes are in the cytosol (Yazaki et al., 1995; Ruelland et al., 2003; Achnine et al., 2005; Dhaubhadel et al., 2008; Hugueney et al., 2009), although members of both families have been reported in other compartments (Halkier and Møller, 1989; Anhalt and Weissenböck, 1992; Ibrahim, 1992; Jones and Vogt, 2001). The localization of the As-SCPL acyltransferase to the vacuole is consistent with predictions for this family (Fraser et al., 2005; Mugford et al., 2009) and with experimental data for the As-SCPL acyltransferase, which was found to be vacuolar (Hause et al., 2002).

An interesting aspect of the regulation of the avenacin biosynthetic pathway has emerged as a part of this work. We found...
that the As-SCPL1 acyltransferase protein, which catalyzes the final step in avenacin biosynthesis, was not detectable in mutants that are blocked in early stages of the pathway (sad1 and sad2 mutants) but was detectable in mutants blocked late in the pathway (sad7 and sad9 mutants; Figure 3; see Supplemental Figure 2 online). As-SCPL1 transcript levels were not altered in any of the mutants tested (see Supplemental Figure 2 online). This suggests that the levels of As-SCPL1 may be subject to post-transcriptional regulation, possibly in response to the presence of an intermediate in the avenacin biosynthetic pathway. This regulatory mechanism seems to be specific to As-SCPL1, since other avenacin biosynthetic enzymes are not affected in the same way. This adds another level of regulation to the tight cotranscriptional regulation that restricts the expression of the avenacin gene cluster to root tip epidermal cells.

The mechanisms by which plant natural product gene clusters form are not understood. It has been speculated that the formation of gene clusters encoding plant natural product biosynthetic pathways is associated with the toxicity of pathway intermediates and is driven by selection to prevent partial pathway inheritance (Takos and Rook, 2012). However, we found that the sad9 mutants do not exhibit any obvious growth defects or other abnormalities that might be expected as a result of the accumulation of a toxic intermediate compound. This is similar in the case of other mutants blocked in this pathway, with the exception of the sad3 and sad4 mutants, which accumulate a toxic form of avenacin that is incompletely glucosylated and that disrupts root development (Mylona et al., 2008). While containment of toxic intermediates may contribute to the selection driving the evolution of plant natural product gene clusters, it seems likely that other factors, such as coinheritance of beneficial combinations of alleles of the different pathway genes and coordinated regulation at the level of chromatin, are also likely to be important (Wegel et al., 2009; Chu et al., 2011).

Previously, we reported that several of the avenacin biosynthetic genes bear signatures of recent and strong selective pressure (Qi et al., 2004, 2006; Mugford et al., 2009; Mugford and Osbourn, 2010), consistent with a change of function since the divergence of these sequences from their relatives in other species. It remains to be established whether the neofunctionalization of these enzymes preceded the clustering of their genes or vice versa. In-depth phylogenetic analysis coupled with comparative genomics for all of these large multigene families will be required to address these questions.

**METHODS**

**Plant Material**

The wild-type and mutant *Avena strigosa* lines used in this study are described by Papadopoulou et al. (1999), Qi et al. (2006), and Mugford et al. (2009). Pathogenicity assays were performed as previously described (Papadopoulou et al., 1999).

**BAC Contig Expansion and Expression Analysis**

The *A. strigosa* BAC contig spanning the Sad1, Sad2, Sad7, and UGT74H5 (Sad10) genes has been previously described (Qi et al., 2006; Mugford et al., 2009; Owatworakit et al., 2012). Overlapping BAC clones were identified by screening the BAC library with single/low-copy number probes from the ends of BAC clones and verified by PCR and restriction mapping. New BAC clones were verified by shotgun sequencing.

The predicted full-length coding sequence of the Sad9 gene was amplified by PCR and cloned into the pET24a vector (Novagen) with the addition of sequence encoding a 6XHis epitope tag to the C terminus of the protein. The His-tagged Sad9 protein was expressed in *Escherichia coli* (BL21) and purified by Ni-afinity chromatography and subsequently by gel filtration chromatography according to standard methods. Purified protein (1 mg) was supplied to Biogenes for the production of antisera in rats. Immunoblot analysis was performed as described by Mugford et al., (2009), and the anti-Sad9 antisera was used at a dilution of 1:500.

RNA was extracted and cDNA synthesized as described previously (Qi et al., 2006).

Quantitative real-time PCR was performed using a Bio-Rad CFX96 real-time system C1000 thermal cycler using SYBR Green JumpStart Taq ReadyMix for quantitative PCR (Sigma-Aldrich). Oligonucleotide sequences used were *EcoR*, 5’-TCCCATCTGGATTGAG-3’ and 5’-TCTCTTGTTGCTGATCG-3’; As-SCPL1, 5’-GCTCAGTCGATAGTCC-3’ and 5’-GATCCATTCGGACATG-3’; and As-MT1, 5’-ACCCCAAGCACATCAAG-3’ and 5’-CGCGGTCTCTAGGATTG-3’.

**Phylogenetic Methods**

Protein sequence alignment was performed using Muscle 3.6 (Edgar 2004) and the neighbor joining phylogeny was constructed using Mega 4.0 (Tamura et al., 2007). The full sequence alignment used to reconstruct the phylogeny can be found in Supplemental Data Set 1 online. Accession numbers for sequences included in the phylogeny are listed below. The abbreviations used are as follows: CNMT, (S)-coaccharine N-methyltransferase; 4OMT, (S)-3-hydroxy-N-methylcoclaurine 4-O-methyltransferase; 6OMT, (S)-norcoaccharine 6-O-methyltransferase; 7OMT, (R,S)-reticuline 7-O-methyltransferase; BANMT, β-Ala N-methyltransferase; BOMT, benzoic acid carboxyl methyltransferase; CaOMT, catechol O-methyltransferase; ChOMT, chavicol O-methyltransferase; CoOMT, caurantine O-methyltransferase; DMXNMT, dimethyllanthine N-methyltransferase; EOMT, (iso)eugenol O-methyltransferase; FOMT, flavonoid 8-O-methyltransferase; MXNMT, 7-methyllanthine N-methyltransferase; PEAAnT, phosphoethanolamine N-methyltransferase; PNMT, putrescine N-methyltransferase; SANT, salicylic acid carboxyl methyltransferase; SOTM, scoulerine 9-O-methyltransferase; and XNMT, xanthoxine methyltransferase.

**sad9 Mutant Identification and Metabolite Analysis**

Previously uncharacterized avenacin-deficient *A. strigosa* mutants (Papadopoulou et al., 1999; Qi et al., 2006) were screened for single nucleotide polymorphisms in the Sad9 gene using the Surveyor mutation detection kit (Transgenomic) according to the manufacturer’s protocol and candidates verified by sequencing. Metabolites were extracted from root tips of 5-d-old seedlings as described previously (Mugford et al., 2009). Fluorescent compounds were analyzed by LC-MS, coupled with fluorescence detection (Mugford et al., 2009). Glic esters of anthranilate and N-methyl anthranilate were quantified by HPLC (also described in Mugford et al., 2009).

**Biochemical Characterization of As-MT1**

Activity assays were performed using 4 μg of purified recombinant As-MT1 protein in 50 mM Tris-HCl, pH 8.0, and 2.5 mM S-adenosyl-Met, in a volume of 200 μL at 30°C. Reactions were started by the addition of the methyl acceptor substrate and stopped after 7 min by the addition of 10 μL of trichloro-acetic acid (TCA). For the determination of kinetic parameters, activity was tested at anthranilate concentrations ranging from 0.05 to 7.5 mM and anthraniloyl-O-Glc from 0.5 to 5 mM, each tested in triplicate.
Concentrations of methylated product were determined by HPLC using a Dionex UltiMate 3000 HPLC with a Luna C18 160 x 4, 60-mm column (Phenomenex) running a gradient of acetonitrile against 0.1% trifluoroacetic acid in water from 10 to 90% over 22 min. The flow rate was constant at 0.5 mL/min, the injection volume was 60 µL, and the detection wavelength used was 320 nm. Data were plotted as a Lineweaver-Burk plot, and Michaelis-Menten kinetic parameters were determined.

CPMV-Based Coexpression

The coding sequences of the As-MT1, As-SCPL1, and As-UGT74H5 cDNAs were amplified by PCR and cloned into the modified pMM1S2NT plasmid (Carlizares et al., 2006) to flank the sequences by the 5' and 3' untranslated regions of CPMV RNA-2 prior to transfer to pBINPLUS. Expression from the cassette was under control of the cauliflower mosaic virus 3S promoter. Constructs were transformed into Agrobacterium tumefaciens LBA4404 and infiltrated into Nicotiana benthamiana leaves as described previously (Carlizares et al., 2006; Mugford et al., 2009). For coexpression experiments, Agrobacterium cultures carrying each of the three constructs were mixed prior to infiltration or were diluted in infiltration medium for expression alone. Leaves were harvested after 6 d and snap-frozen in liquid nitrogen or in Figure 2 are as follows: Thalictrum flavum BMT, AAU20765; Coptis japonica 6OMT, BAB08004; Papaver somniferum 6OMT, AAQ01669; P. somniferum 4OMT, AAP45313; T. flavum 4OMT, AAU20768; C. japonica 4OMT, BAB08005; C. japonica CoOMT, BAC22084; A. strigosus MT1, JQ071450; Mentha x piperita FOMT, AY337459; Ocimum basilicum EOMT, AAL30424; O. basilicum ChOMT, AAL30423; P. somniferum 7OMT, AAQ01668; Limonium latifolium BANMT, AAP03058; T. flavum SMT, AAU20770; C. japonica SMT, BAA06192; Anmmis majus COMT, AAC24095; R. graveolens ANMT, DQ849332; P. somniferum CaOMT, AAQ01670; O. basilicum COMT1, AAD38189; O. basilicum COMT2, AF159418; Clarkia breweri EOMT, AAC01533; C. breweri COMT, O23760; Atropa belladonna PNMT, BAA82264; Datura stramonium PNMT, CAE47481; Solanum lycopersicum PEANMT, AAGS5894; Coffea arabica DMXNMT1, BAC75663; C. arabica MXNMT, BAB39216; C. arabica XNMT, BAB39215; C. japonica PNMT, BAB12278; Arttinhimum majus BOMT, AAF88284; C. breweri SAMT, AAF00108; Mesorozobium loti CFAPS, BAB53730; Arabidopsis At4g33110, AAM57626; Arabidopsis At4g33120, ORF NP 195038; P. somniferum CNMT, AAP45316; T. flavum CNMT, AAU20766; and C. japonica CNMT, BAB71802.

Supplemental Data

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

Supplemental Figure 1. Wild-Type Oat Root Sections Probed with Preimmune Sera.

Supplemental Figure 2. Identification of Des-Methyl Avenacin A-1 and Anthraniloyl-Glc by LC-MS.

Supplemental Figure 3. 1H-NMR Spectra of the Avenacin Triterpene.

Supplemental Figure 4. Acylation of Avenacins Is Important for Biological Activity.

Supplemental Figure 5. As-SCPL1 Expression Is Reduced or Abolished in Oat Mutants Blocked in the Early Steps in the Synthesis of the Avenacin Triterpene.

Supplemental Figure 6. LC-MS Analysis of Fluorescent Compounds from N. benthamiana Plants Coexpressing As-MT1 and As-UGT74H5.

Supplemental Figure 7. Fluorescent Compounds Accumulate in the Vacuole of sad7 Mutants.

Supplemental Table 1. Quantification of Avenacins and Des-Methyl Avenacin A-1 in Wild-Type and Mutant Oat Root Tips.

Supplemental Table 2. Quantification of Fluorescent Compounds in N. benthamiana Leaves Coexpressing Avenacin Biosynthetic Enzymes.

Supplemental Table 3. Subcellular Localization of the As-MT1, As-UGT74H5, and As-SCPL1 Proteins.

Supplemental Data Set 1. Protein Sequence Alignment Used in Phylogenetic Analysis.

ACKNOWLEDGMENTS

We thank Kim Findlay and Sue Bunnewell for immunogold labeling and transmission electron microscopy, Grant Calder for confocal microscopy, and Ian Colquhoun for NMR. This project was funded by the Biotechnology and Biological Sciences Research Council, UK (Grant BB/
E009912/1 to S.T.M.), by a joint Engineering and Physical Sciences Research Council/National Science Foundation award to A.O. and S.J.R. as part of the Syntegron Consortium (EP/H019154/1), and by the University of Glasgow work placement program (S.H.). R.M., L.H., G.P.L., and A.O. are supported by the Biotechnology and Biological Sciences Research Council Institute Strategic Program Grant “Understanding and Exploiting Plant and Microbial Secondary Metabolism” (BB/J004561/1) and the John Innes Foundation.

AUTHOR CONTRIBUTIONS

Received February 7, 2013; revised February 25, 2013; accepted March 5, 2013; published March 29, 2013.

REFERENCES


Correction


Figure 1 has been corrected to read:

![Figure 1. Structures of Oat Root Avenacins.](image)

The C-21 acyl group consists of either N-methyl anthranilate or benzoate, depending on the R1 group. The major avenacin form found in oat roots is avenacin A-1, which is acylated with N-methyl anthranilate (Crombie et al., 1984).
Modularity of Plant Metabolic Gene Clusters: A Trio of Linked Genes That Are Collectively Required for Acylation of Triterpenes in Oat
Sam T. Mugford, Thomas Louveau, Rachel Melton, Xiaoquan Qi, Saleha Bakht, Lionel Hill, Tetsu Tsurushima, Suvi Honkanen, Susan J. Rosser, George P. Lomonossoff and Anne Osbourn

Plant Cell 2013;25;1078-1092; originally published online March 26, 2013;
DOI 10.1105/tpc.113.110551

This information is current as of December 31, 2017

Supplemental Data /content/suppl/2013/03/07/tpc.113.110551.DC1.html
References This article cites 70 articles, 30 of which can be accessed free at:
/content/25/3/1078.full.html#ref-list-1
eTOCs Sign up for eTOCs at:
http://www.plantcell.org/cgi/alerts/ctmain
CiteTrack Alerts Sign up for CiteTrack Alerts at:
http://www.plantcell.org/cgi/alerts/ctmain
Subscription Information Subscription Information for The Plant Cell and Plant Physiology is available at:
http://www.aspb.org/publications/subscriptions.cfm
Errata An erratum has been published regarding this article. It is appended to this PDF and can also be accessed at:
/content/25/9/3632.full.pdf

© American Society of Plant Biologists
ADVANCING THE SCIENCE OF PLANT BIOLOGY