***New Phytologist* Supporting Information**

Article title: Towards take-all control: A C-21β oxidase required for acylation of triterpene defense compounds in oat

Authors: Aymeric Leveau, James Reed, Xue Qiao, Michael J. Stephenson, Sam Mugford, Rachel E. Melton, Jenni C. Rant, Robert Vickerstaff, Tim Langdon, and Anne Osbourn.

The following Supporting Information is available for this article:

Fig. S1 Electrospray ionization (ESI) mass spectra for unacylated avenacins detected in roots of *A. strigosa* avenacin-deficient mutants shown in Fig. 1c.
Fig. S2 GC-MS total ion chromatograms (TIC) from analysis of extracts of *N. benthamiana* leaves expressing AstHMGR and AsbAS1, or AstHMGR, AsbAS1 and AsCYP51H10 with candidate CYPs.
Fig. S3 Electron-impact mass spectra for 12,13β-epoxy-16β-hydroxy-β-amyrin (EpHβA) and putative 12,13β-epoxy-16,21β-hydroxy-β-amyrin (EpdiHβA).
Fig. S4 NMR assignment for 12,13β-epoxy,16β,21β-dihydroxy-β-amyrin-3β-O-L-arabinose.
Fig. S5 Electron-impact mass spectra for the unknown product coeluting with 21-hydroxy-β-amyrin.
Fig. S6 Identification of additional *AsCYP72A475* (*sad6*) mutants by LC-MS metabolic profiling.
Fig. S7 Schematic of AsCYP72A475 and mutation sites in *sad6* mutants.

Table S1 Oligonucleotide primer sequences used for Gateway cloning of coding sequences of the candidate CYP genes.
Table S2 Oligonucleotide primer sequences used for sequencing of the genomic DNA region encoding for the *Sad6* gene.
Table S3 Oligonucleotide primer sequences used for RT-PCR expression profiling of the candidate CYPs from *A. strigosa*.

Table S4 Oligonucleotide primer sequences Primers used for quantitative PCR expression profiling of *A. strigosa* CYPs.

Methods S1 Metabolite extraction, LC-MS-MS and LC-MS-Fluorescence analysis of oat roots.

Methods S2 Metabolite extraction and LC-MS-CAD analysis of extracts of *N. benthamiana* leaves

Methods S3 Purification and structural elucidation of 12,13-epoxy,16,21-dihydroxy-β-amyrin-3-*O*-L-arabinose.

Methods S4 Structural analysis of AsCYP72A475.

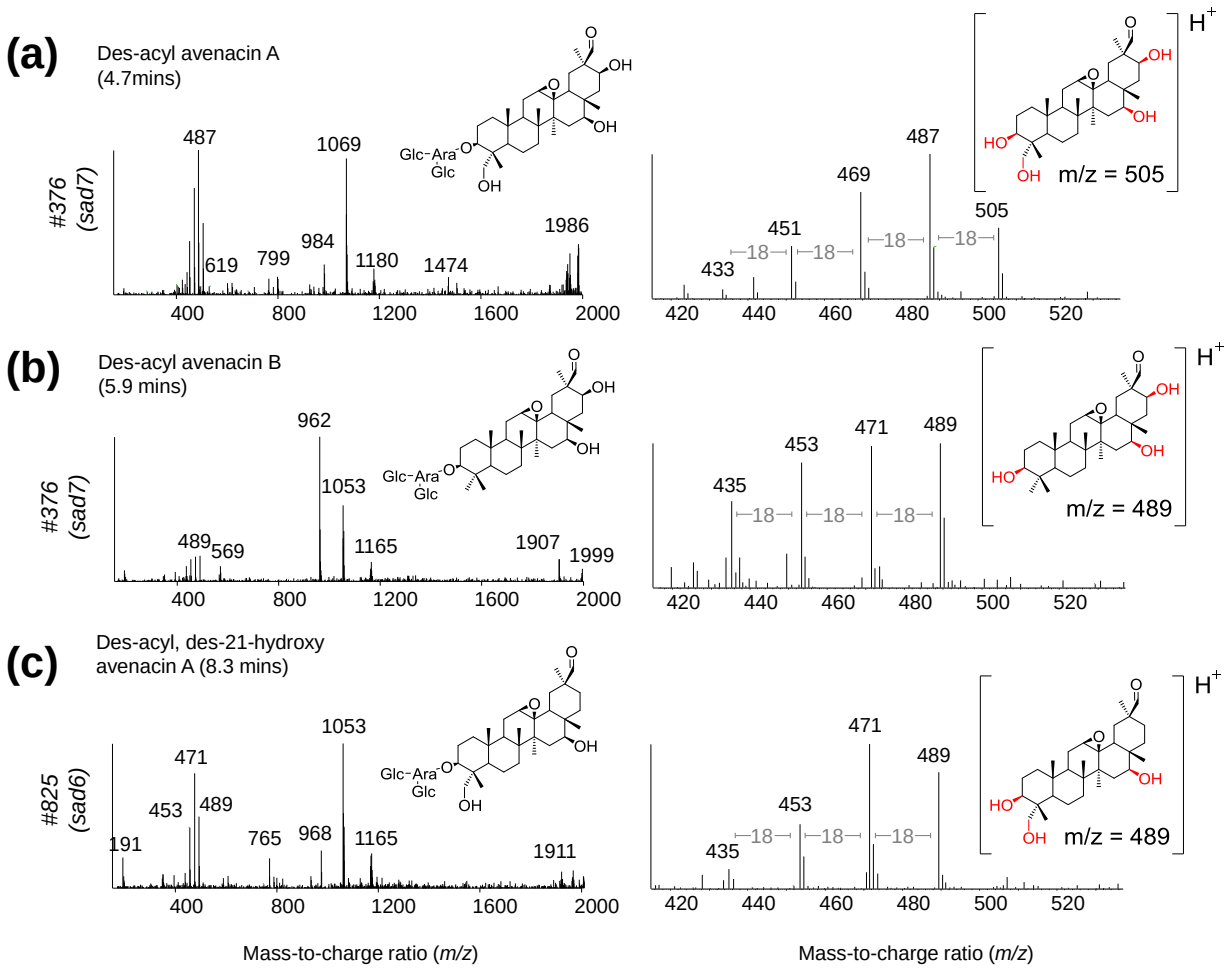


Fig. S1 Mass spectra for unacylated avenacins detected in roots of *A. strigosa* avenacin-deficient mutants shown in Fig. 1c. **a)** Des-acyl-avenacin A, as found in acyltransferase-deficient *sad7* mutants. **b)** Des-acyl-avenacin B, as found in *sad7* mutants. **c)** New compound detected in *sad6* mutants. Left, full mass spectra; right, ions within these spectra that likely represent in-chamber fragmentation of the aglycone. The aglycone shows characteristic losses of *m/z* 18 (loss of the hydroxyl groups as water, indicated in grey). The spectrum for the compound with the 8.3 min retention time in root extracts of *sad6* mutants is consistent with an avenacin-like molecule that lacks the acyl group and contains three hydroxyl groups. Together these data suggest that the C-21β hydroxyl group is most likely absent. Hence the compound at 8.3 mins was assigned as des-acyl, des-21-hydroxy avenacin A. Hydroxyl groups are indicated in red.

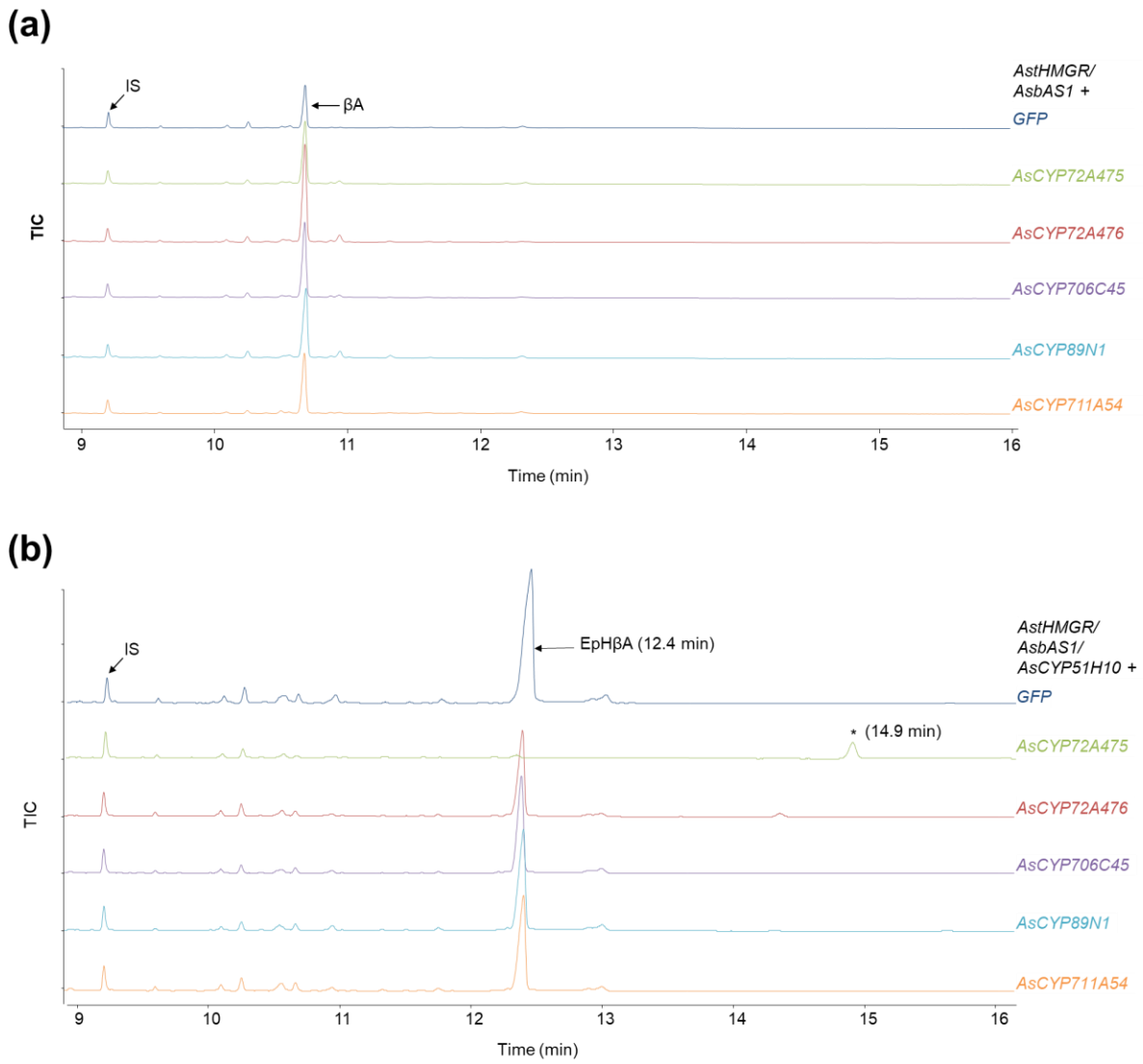
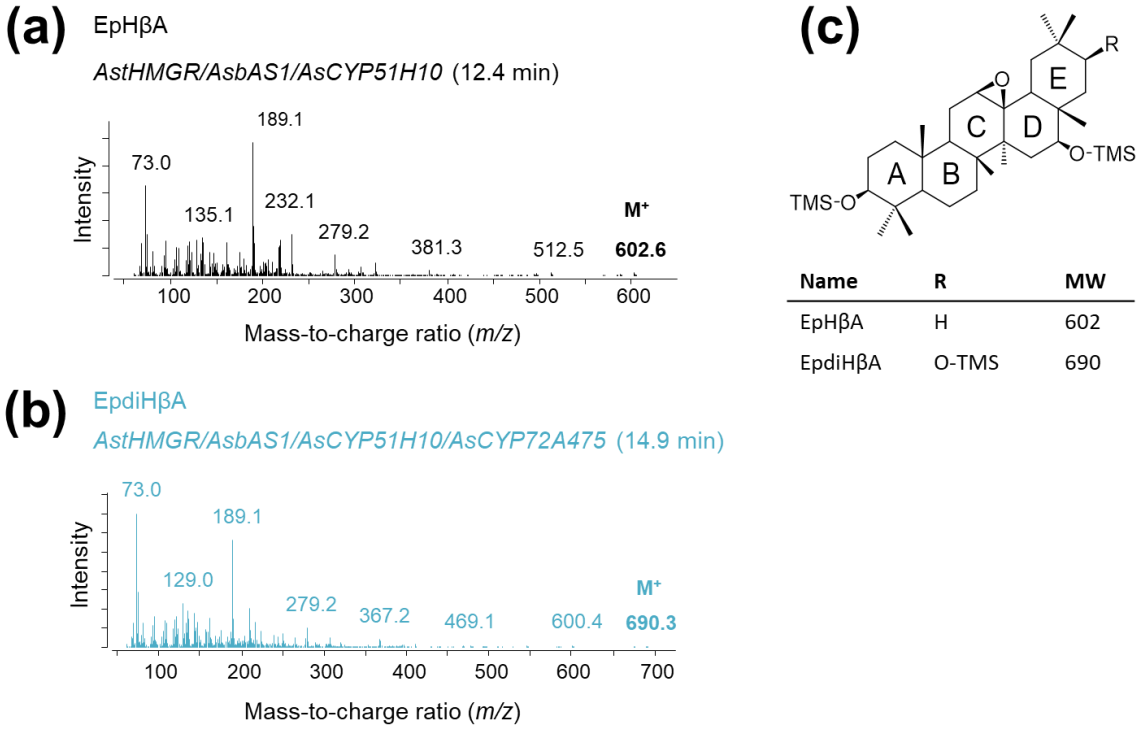


Fig. S2 GC-MS total ion chromatograms (TIC) from analysis of extracts of *N. benthamiana* leaves expressing **a)** AstHMGR and AsbAS1, or **b)** AstHMGR, AsbAS1 and AsCYP51H10 with candidate CYPs. Activity of candidate CYPs towards β-amyrin (βA) was not detected. However, conversion of the β-amyrin scaffold to 12,13β-epoxy, 16β-hydroxy-β-amyrin (EpHβA) by CYP51H10, resulted in a new product at 14.9 min in extracts from leaves expressing AsCYP72A475. The new compound was assigned an identity of 12,13β-epoxy, 16β,21β-dihydroxy-β-amyrin (EpdiHβA). Mass spectra for EpHβA and EpdiHβA are shown in Fig. S3. Chromatograms were normalised relative to the coprostanol internal standard peak area (IS).



Name	R	MW
EpHβA	H	602
EpdiHβA	O-TMS	690

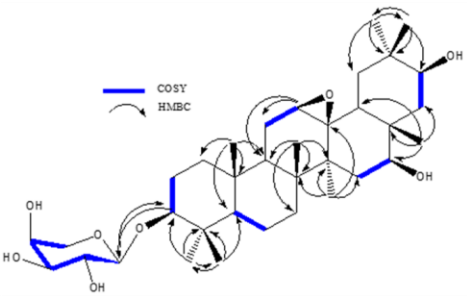
(d)

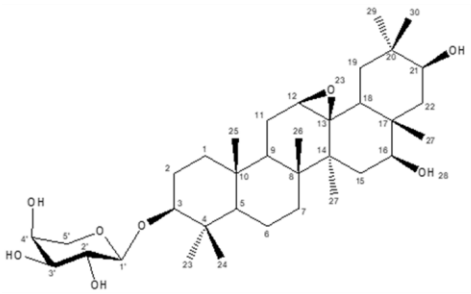
Predicted fragment loss	12,13β-Epoxy-16β-hydroxy-β-amyrin	12,13β-Epoxy-16,21β-dihydroxy-β-amyrin
None	602	690
CH_3•	587	675
TMSiOH•	512	600
2 x TMSiOH•	422	510
TMSiOH• / CH_3•	497	585
ABC*-rings	322	410
ABC*-rings / TMSiOH	232	320
ABC*/C28 moiety	307	397
ABC*/C28 moiety /TMSiOH	217	307
ABC*/C28 moiety / 2 x TMSiOH	-	217
C*DE-rings	279	279
C*DE-rings - TMSiOH	189	189

C* Represents a partial C ring

Fig. S3 Electron-impact mass spectra for 12,13β-epoxy-16β-hydroxy-β-amyrin (EpHβA) and putative 12,13β-epoxy-16,21β-hydroxy-β-amyrin (EpdiHβA). **a), b)** Mass spectra for products shown in Fig. 3a and Fig. S2b. The putative molecular ions (M^+) are labelled. **c)** Structure of the proposed trimethylsilylated (TMS) EpHβA and EpdiHβA. **d)** Characteristic ions found in these mass spectra and their expected corresponding fragments. The ring assignments are as shown in **c)**.

Carbon numbering scheme and selected COSY and HMBC correlations





Carbon number	^{13}C δ	^{1}H δ	Carbon number	^{13}C δ	^{1}H δ
1'	107.16	4.25 (1H, d, *J*=6.7)	22	39.5596	1.10 (1H, dd, *J*= 13.7, 4.4) 1.09 (1H, m)
3	90.47	3.10 (1H, dd, *J*=11.7, 4.5)	1	39.1961	1.53 (1H, m) 0.95 (1H, m)
3'	74.33	3.50 (1H, m)	10	38.0548	/
21	73.85	3.47 (1H, m)	20	36.8509	/
2'	72.82	3.56 (1H, m)	7	35.3702	1.47 (1H, m) 1.18 (1H, m)
4'	69.52	3.79 (1H, m)	15	34.7066	1.86 (1H, m) 1.37 (1H, m)
13	68.03	/	29	29.7707	0.97 (3H, s)
16	67.78	4.11 (1H, dd, *J*= 11.0, 4.8)	23	28.4099	1.02 (3H, s)
5'	66.40	3.82 (1H, m) 3.50 (1H, m)	2	26.9553	1.84 (1H, m) 1.66 (1H, m)
5	57.05	0.74 (1H, m)	27	23.895	1.26 (3H, s)
12	55.56	2.79 (1H, d, *J*=3.4)	11	22.6372	1.82 (2H, m)
18	50.51	1.14 (1H, m)	28	22.5442	1.00 (3H, s)
9	46.96	1.41 (1H, m)	26	21.0644	1.11 (3H, s)
14	43.32	/	6	18.7245	1.52 (2H, m)
8	42.68	/	30	17.2815	0.77 (3H, s)
19	42.61	1.83 (1H, m) 1.27 (1H, m)	24	16.7204	0.83 (3H, s)
17	42.13	/	25	15.6094	0.90 (3H, s)
4	40.22	/			

s = singlet
d = doublet
dd = doublet of doublets
t = triplet
m = multiplet

Fig. S4 NMR assignment for 12,13β-epoxy,16β,21β-dihydroxy-β-amyrin-3β-*O*-L-arabinose. The purified product from *N. benthamiana* leaves co-expressing AstHMGR, AsbAS1, AsCYP51H10, AsCYP72A475 and AsAAT1 was subjected to NMR analysis. ^{1}H and ^{13}C spectra were fully assigned using a combination of DEPT-edited-HSQC, HMBC, COSY, and DEPT-135 experiments. The C-21 OH was assigned by a 2D-NOESY experiment as β due to NOE cross-peaks observed between the C-21 H and the C-16 H, and between the C-21 H and the C-29 H_3. Spectra were recorded in MeOH-*d*4 and the chemical shifts referenced to an internal tetramethylsilane (TMS) standard.

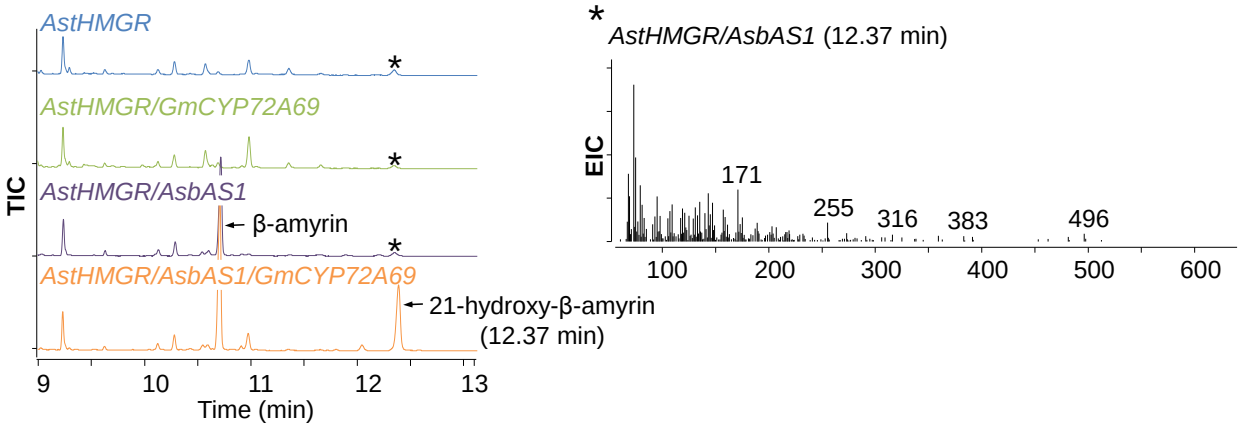
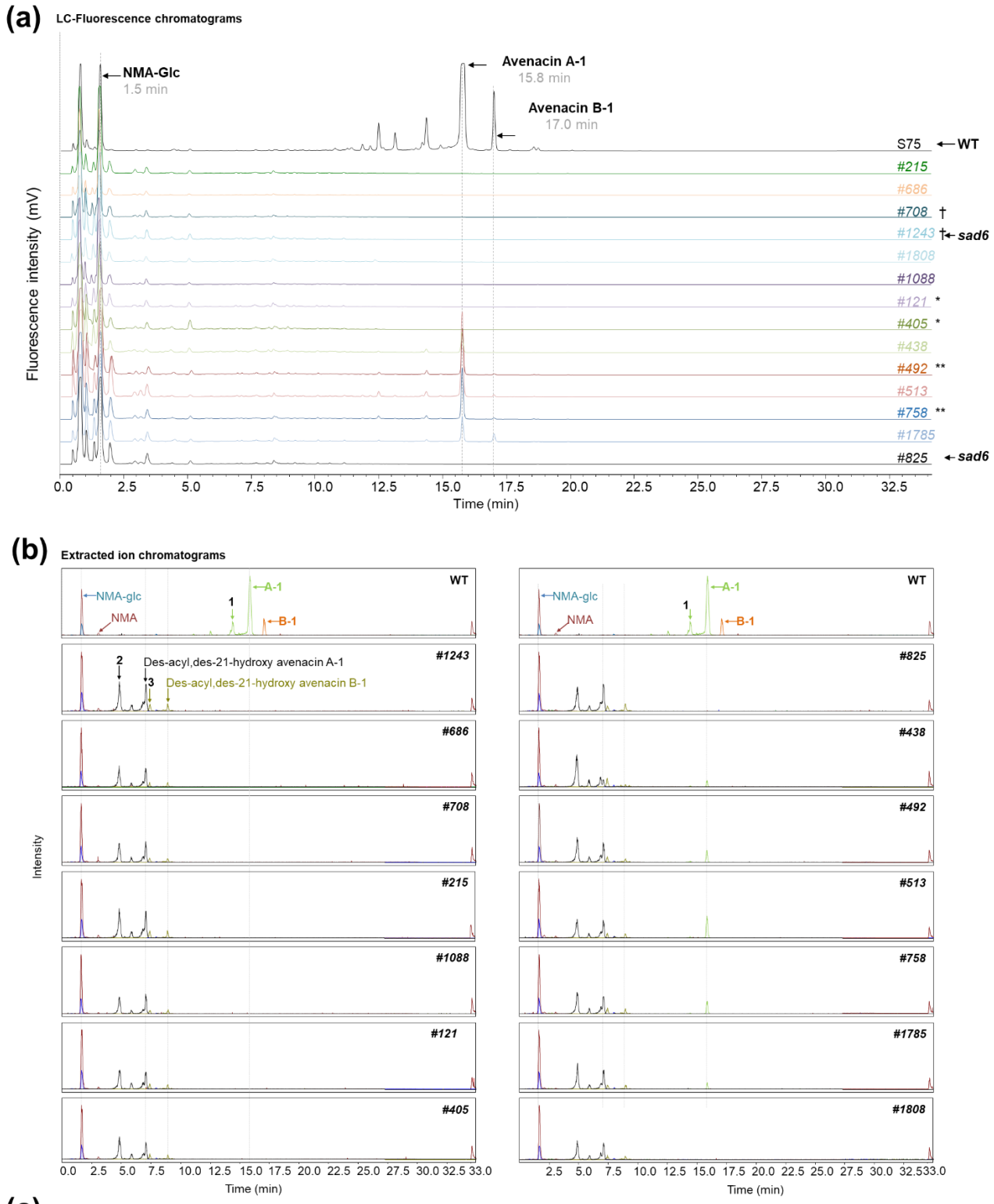


Fig. S5 Electron-impact mass spectra for the unknown product coeluting with 21-hydroxy-β-amyrin. Left, GC-MS total ion chromatograms (TIC). An unknown product (*) coeluted with the putative 21-hydroxy-β-amyrin. The mass spectrum for this product is shown on the right.



(c)

Compound	Event	Specific Ion	Retention time (min)
Avenacin A-1 (A-1) $[M/H]^-$	NEG	*m/z 1092*	15.9
Avenacin B-1 (B-1) $[M/H]^-$	NEG	*m/z 1076*	17.1
Des-acyl,des-21-hydroxy avenacin A-1 $[M/COO]^-$	NEG	*m/z 989*	7.1
Des-acyl,des-21-hydroxy avenacin B-1 $[M/COO]^-$	NEG	*m/z 973*	8.9
N-Methyl anthraniloyl-*O*-glucoside (NMA-glc) $[M/H]^+$	POS	*m/z 314*	1.6
N-Methyl anthranilate (NMA) $[M/H]^+$	POS	*m/z 152*	3.0

Fig. S6 Identification of additional *AsCYP72A475* (*sad6*) mutants by LC-MS metabolic profiling. **a)** LC-fluorescence (Ex:353 nm, Em:441 nm) analysis of root extracts of the *sad6* mutants. Avenacins were not detectable in the two original *sad6* mutants (***sad6***), the new *sad6* mutants harbouring premature stop codons (#215, #686, #708, #1243 *and* #1808), the #1088 splice site mutant, the #815 mutant (likely indel), or the two mutants carrying a G493R amino acid substitution mutation (#121 and #405). Conversely, small amounts of avenacin A-1 could still be detected in the five remaining *sad6* mutants (#438, #492, #513, #758 and #1785), suggesting that these mutations do not completely disrupt the function of AsCYP72A475. A list of the mutant lines and the identified *AsCYP72A475* mutations are provided in Fig. S7a. *N*-methyl anthraniloyl-*O*-Glucose (NMA-glc) is detectable in the wild type and all *sad6* mutants. **b)** LC-MS analysis of root extracts of *sad6* mutants. Extracted ion chromatograms (EICs) show that all these mutants are impaired in avenacin production and accumulate des-acyl, des-21-hydroxy avenacin A-1 and des-acyl, des-21-hydroxy avenacin B-1 instead. Peaks 1, 2 and 3 are likely degradation products of avenacin A-1, des-acyl, des-21-hydroxy avenacin A-1, and des-acyl, des-21-hydroxy avenacin B-1 respectively. Degradation likely results from the high reactivity of the 12,13β-epoxy group present on the β-amyrin scaffold of these compounds. **c)** Summary of the parameters used for the LC-MS EIC analysis for each of the compounds monitored in **b)**.

(a)

Mutant	Mutation event	Predicted amino acid change	Region of protein
Premature termination of translation:			
#215	G736A	W216 stop	—
#686	G3143A	W 447 stop	—
#708	G2772A	W 356 stop †	—
#1243	G2772A	W356 stop †	—
#1808	G1013A	W195 stop	—
Splicing error:			
#1088	G2501A	—	—
Amino acid substitution:			
#121	G3279A	G493R *	αL helix
#405	G3279A	G493R *	αL helix
#438	C3424T	P428L	β2-2 sheet
#492	G788A	G150D **	SRS1
#513	G3576A	G479R	Haem binding loop
#758	G788A	G150D **	SRS1
Indel mutation:			
#1785	858_872del	256_260del	SRS2
#825	indel?	?	?

(b)

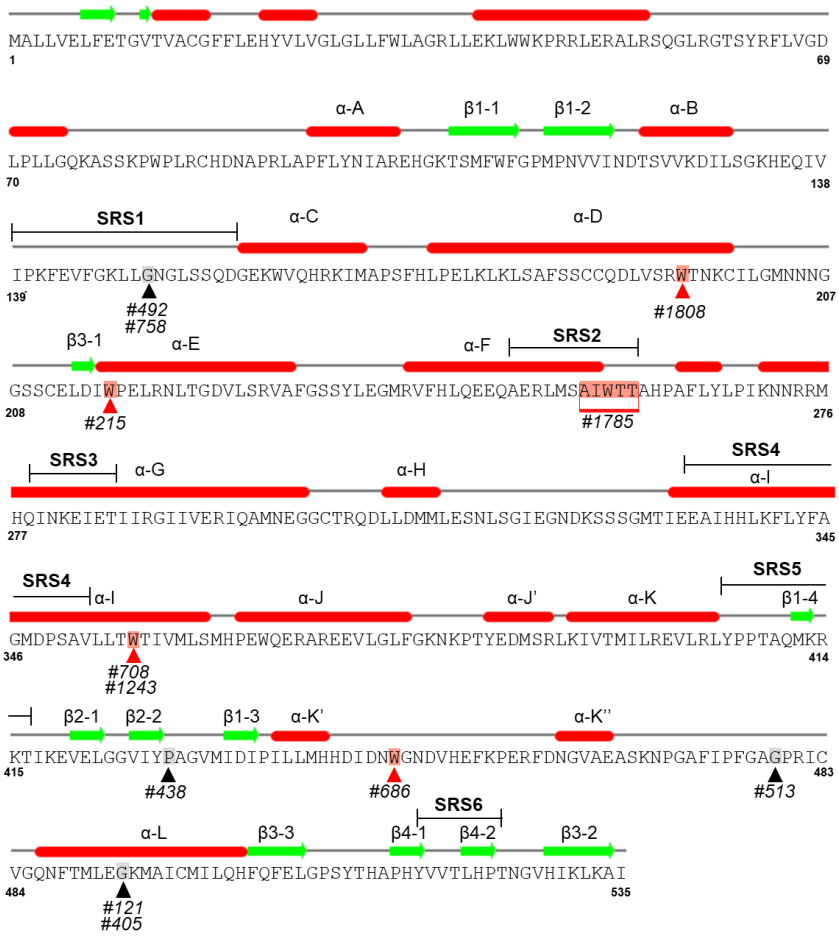


Fig. S7 AsCYP72A475 and mutation sites in *sad6* mutants.

a) Mutant lines and the corresponding mutations identified in the *AsCYP72A475* gene. In some cases, the same mutation was found to occur in different lines (denoted with *, ** and †). However, these represent independent mutation events as the lines are derived from different M2 families. **b)** AsCYP72A475 primary structure with predicted α-helices (red), β-sheets (green) and substrate recognition sites (SRS) 1-6 shown (Werck-Reichhart & Feyereisen, 2000) (see Methods S4). Triangles below the sequence indicate the sites of *sad6* predicted amino acid mutations: red, nonsense mutations; black, substitutions.

Table S1 Oligonucleotide primer sequences used for Gateway cloning of coding sequences of candidate CYP genes

Primer	Sequence (5'-3')
CYP71E22-GW-F	AAAAAGCAGGCTTCATGGTGATCCTGCACCCCTCGC
CYP71E22-GW-R	AGAAAGCTGGGTCCTAGGCGCGCGGCGGCACGTATAG
CYP72A475-GW-F	AAAAAGCAGGCTTCATGGCACTGCTTGTAGAGTTGTTC
CYP72A475-GW-R	AGAAAGCTGGGTCTCATATAGCCTTGAGCTTAATATG
CYP72A476-GW-F	AAAAAGCAGGCTTCATGAGTAGTTTTGAGATGGAATTC
CYP72A476-GW-R	AGAAAGCTGGGTCMGATGGCCTTGAGAGTAATCTGAGC
CYP88A75-GW-F	AAAAAGCAGGCTTCATGTGGTGGCCGTGGTGGGCGG
CYP88A75-GW-R	AGAAAGCTGGGTCTCAGTGTCCACCGGAGCCGGAGG
CYP89N1-GW-F	AAAAAGCAGGCTTCATGGAGGGGCTCACCATCCTCTTG
CYP89N1-GW-R	AGAAAGCTGGGTCMCTGCTTCATGTCCCTTGGGATGATG
CYP706C45-GW-F	AAAAAGCAGGCTTCATGCCACATACGTTAACGTTTTCC
CYP706C45-GW-R	AGAAAGCTGGGTCTTACTCAGACGAATAATAGAGTTC
CYP711A54-GW-F	AAAAAGCAGGCTTCATGGGCCGCCTGGCAGAAGCACTG
CYP711A54-GW-R	AGAAAGCTGGGTCTCATATGTTCCGCTCGATGACTTGGAG
attB1-adapter	GGGGACAAGTTTGTACAAAAAAGCAGGCTTC
attB2-adapter	GGGGACCACTTTGTACAAGAAAGCTGGGTC

Table S2 Oligonucleotide primer sequences used for sequencing of the genomic DNA region encoding for the *AsCYP72A475* gene

Primer	Sequence (5'-3')
For-6	AGGAAGTTGAGAATCGAGCTGAG
For-1	ATGGCACTGCTTGTAGAGTTGTTC
For-2	GAATATTATTTGTTACTCCGACAC
For-3	GATATCAATATATGCTCTCTGTAC
For-4	CATATGAAGACATGAGCCGGCTG
Rev-1	GGATGCACTTATTCGTCCATCTGC
Rev-2	CGAGGTAGACGATCATGATGTAGC
Rev-3	CGCGAAGAATCATGGTAACCTGC
Rev-4	TCATATAGCCTTGAGCTTAATATG

Table S3 Oligonucleotide primer sequences used for RT-PCR expression profiling of candidate CYPs from *A. strigosa*.

Amplicon	Primer	Sequence (5'-3')
Candidate CYPs	CYP72A475-RT-PCR-F	CCGAGGCCTCAAAGAATCCA
	CYP72A475-RT-PCR-R	TGACCCATTATTACACTGTTGCA
	CYP72A476-RT-PCR-F	TTCCAAGGCGCATATCTCCC
	CYP72A476-RT-PCR-R	TCAGATCCCGATTCATCCGC
	CYP706C45-RT-PCR-F	CTCGAGTTCAACCCGGAGAG
	CYP706C45-RT-PCR-R	AATTCAAGTTCATGCCGCCT
	CYP71E22-RT-PCR-F	GGTCAATGGCGCTGATGAAC
	CYP71E22-RT-PCR-R	TTTATTTGGGAACGACGGGC
	CYP89N1-RT-PCR-F	CACACTGCATGCCGAGTACT
	CYP89N1-RT-PCR-R	TGAATTAAGAGGACCGGGCC
	CYP88A75-RT-PCR-F	GACATGATGGACCGGCTGAT
	CYP88A75-RT-PCR-R	AAAGTGTCTCGAGTCGCCTG
	CYP711A54-RT-PCR-F	GCGATACGGGCCAATCTACA
	CYP711A54-RT-PCR-R	CGGCCCTGTTGATGTAAGGT
Controls	GAPDH-RT-PCR-F	GGTGGTCATTTCAGCCCCTA
	GAPDH-RT-PCR-R	CTCCCACCTCTCCAGTCCTT
	AsbAS1-RT-PCR-F	AACTACGCCAACTACCGCAA
	AsbAS1-RT-PCR-R	ACGATTAAACCGCGGGACTT
	CYP51H10-RT-PCR-F	ACACACTTACAAGGACCCCG
	CYP51H10-RT-PCR-R	CGCACACAAAGATACATCCGT

Table S4 Oligonucleotide primer sequences used for qPCR expression profiling.

Primer	Sequence (5'-3')
EF1-α-qPCR-F	TCCCCATCTCTGGATTTGAG
EF1-α-qPCR-R	TCTCTTGGGCTCGTTGATCT
AsbAS1-qPCR-F	TGCGGAATTCACAAAGAACA
AsbAS1-qPCR-R	GCTTGGCTTCTGTCGGAATA
CYP51H10-qPCR-F	TCGACAGGAAGTGGAGGAGT
CYP51H10-qPCR-R	ATCTCGGACCTCACTTCCAA
CYP72A475-qPCR-F	GGAAGCTGCTAGGTAATGG
CYP72A475-qPCR-R	CCGACAACTTGAGCTTGA
CYP72A476-qPCR-F	GCATTCAACGAGATGCTC
CYP72A476-qPCR-R	AAATTGTTCTTGATGGCCTG
CYP94D65-qPCR-F	GTCATTACCAGGTTCGTGT
CYP94D65-qPCR-R	GATGATGTTGTCGGCGTA

Methods S1. Metabolite analysis of wild type and mutant oat lines

Metabolite extractions

Root tips from 5-day-old *A. strigosa* seedlings were harvested and ground with tungsten beads in 1 ml of 80% methanol, using a homogenizer with a 24 Hz frequency (2010 Geno/Grinder, SPEX SamplePrep, Stanmore, UK). Ground samples were incubated at room temperature for 30 min with shaking at 1400 rpm (Thermomixer Comfort, Eppendorf). Samples were centrifuged for 10 min at 12,000 g and 0.8 ml of the supernatant was partitioned twice with 0.4 ml of hexane. The aqueous phase was dried in vacuo (EZ-2 Series Evaporator, Genevac, Ipswich, UK), and the dried material resuspended in 170 µL of 50% MeOH and subsequently filtered (0.2 µm, Spin-X, Costar, Sigma-Aldrich, Gillingham, UK). Filtered samples were transferred to glass vials before analysis.

LC-MS-MS analysis

For LC-MS-MS analysis, metabolites were extracted and analysed as described by Mugford et al. (2009).

LC-MS-Fluorescence analysis

Analyses were performed using a Prominence HPLC system, RF-20Axs fluorescence detector, single quadrupole mass spectrometer LC-MS-2020 (Shimadzu, Milton Keynes, UK). The method was designed with a flow rate of 0.3 ml per min and root extracts were separated on a Kinetex column 2.6 µm XB-C18 100 Å, 50 x 2.1 mm (Phenomenex, Macclesfield, UK) column, using the following solvent system: Solvent A: [H_2O + 0.1% formic acid (FA)], Solvent B: [acetonitrile (CH_3CN) + 0.1% FA]. The mobile phase was set to 20% solvent B for the first 3 min, followed by a gradient from 20% to 60% across 25 min. The gradient was raised from 60% to 100% for 2 min and held at 100% for the final 3 min. The fluorescence detector was set to excitation at 353 nm and emission at 441 nm. The mass spectrometer utilised dual ESI/APCI ionization, DL temp 250 °C, nebulizer gas flow rate 15 L per min, heat block temperature 400 °C, spray voltage positive (4.5 kV), Negative (-3.5 kV).

Methods S2. Analysis of *N. benthamiana* leaf extracts

Leaves were harvested 6 days post-infiltration and freeze-dried. Freeze-dried material was ground using tungsten beads, in a homogenizer with a 24 Hz frequency (2010 Geno/Grinder, SPEX SamplePrep) before adding 1 ml of 80% methanol. Extractions were performed at 42°C for 30 min with shaking at 1400 rpm (Thermomixer Comfort, Eppendorf). Samples were then centrifuged for 10 min at 12,000 g and 0.8 ml of supernatant was partitioned twice with 0.4 ml of hexane to remove lipids. The aqueous phase was dried in vacuo (EZ-2 Series Evaporator, Genevac). Dried material was resuspended in 170 µL of 50% MeOH and subsequently filtered (0.2 µm, Spin-X, Costar). Filtered samples were then transferred to glass vials for LC-MS-CAD analysis.

LC-MS-CAD analysis was performed using a Prominence HPLC system, with single quadrupole mass spectrometer LC-MS-2020 (Shimadzu) and Corona Veo RS Charged Aerosol Detector (Thermo Scientific, UK). The column, solvent systems and flow rate were as described for the LC-MS-Fluorescence of oat root extracts (above). The mobile phase was set to 25% solvent B for the first 1.5 min, followed by a gradient from 25% to 60% for 19.5 min. The gradient was raised from 60% to 100% across 0.5 min and held at 100% for 2 min. CAD parameters included collection rate (10 Hz), filter constant (3.6 s), evaporator temperature (35 °C), ion trap voltage (20.5 V). MS parameters were as described for LC-MS-Fluorescence of oat root extracts (above).

Methods S3. Purification and structural elucidation of 12,13-epoxy,16,21-dihydroxy-β-amyrin-3-*O*-L-arabinose

A total of 300 *N. benthamiana* plants were vacuum-infiltrated with *Agrobacterium* strains as described by Reed et al. (2017). Infiltrated leaves were collected 7 days post infiltration and lyophilized. A total of 98.4 g of dried leaves was then reduced to a coarse powder by adding 1:5 vol:vol of quartz sand (0.3-0.9 mm) in a 2 L screw cap plastic container and shaking it for a couple of minutes. The powder mixture was loaded into 120 ml extraction cells, as described previously (Reed *et al.*, 2017) and extracted using a SpeedExtractor E-914 (Büchi, Oldham, UK) with one cycle at 90°C and 130 bar pressure and hexane as solvent (to the majority of apolar compounds and chlorophyll), followed by four cycles at 90°C and 130 bar pressure with ethyl acetate. Cycle 1 had zero hold time, and cycles 2 to 5 had 5 min hold times. The run finished with a 2 min

solvent flush and a 6 min N_2 flush. The extraction procedure was repeated until all the powdered leaf material was processed. All ethyl acetate fractions were combined and concentrated to dryness by rotary evaporation under vacuum prior to resuspending in 800 ml of 80% MeOH and successive partitioning three times with 500 ml of hexane to remove unwanted chlorophyll and apolar compounds. Diatomaceous earth (Celite, Sigma-Aldrich) (1-2 g) was then added to the methanol fraction and concentrated to dryness by rotary evaporation under vacuum. The crude solid was reduced to powder and subjected to normal phase flash chromatography using a silica column (SNAP KP-SIL 100 g, Biotage®, Hengoed, UK) coupled to a DCM/MeOH mobile phase gradient (4% to 40% MeOH over 15 column volumes). Collected fractions were analysed using thin layer chromatography (TLC) (with a 9:1 DCM:MeOH mobile phase) and those containing the compound of interest were pooled, mixed with a gram of Celite and concentrated to dryness by rotary evaporation under vacuum. The solid was then reduced to powder and subjected to reverse phase flash chromatography using a 12 gram C18 silica column (SNAP Ultra C18, Biotage®) and a H_2O: MeOH mobile phase gradient (70%-95% MeOH on 40 CV). The purest fractions were selected based on TLC analysis and combined prior to concentrating by rotary evaporation under vacuum. After complete evaporation of the MeOH, the compound of interest appeared insoluble in water and formed a white flake-like precipitate that was collected by centrifugation (5 min at 3200g). The solid was then lyophilized to obtain 81.4 mg of powder. After calculation, the yield of product recovered was 0.82 mg/g leaf dry weight.

NMR spectra were recorded in Fourier transform mode at a nominal frequency of 400 MHz for ^{1}H NMR, and 100 MHz for ^{13}C NMR, using deuterated methanol (MeOH-d4). Chemical shifts were recorded in ppm and referenced to an internal TMS standard. Multiplicities are described as: s = singlet, d = doublet, dd = doublet of doublets; coupling constants are reported in Hertz as observed and not corrected for second order effects. Where signals overlap ^{1}H δ is reported as the centre of the respective HSQC crosspeak (see Fig. S4).

Methods S4. Structural analysis of AsCYP72A475

The primary sequence of AsCYP72A475 was submitted to the I-TASSER (Zhang, 2008) server for homology modelling. The secondary structure of the top-scoring I-TASSER model was visualised using Jalview (Waterhouse *et al.*, 2009). Annotation of the secondary structure was based on comparison to the crystal structure of the rabbit CYP4B1 (PDB 5T6Q) (Hsu *et al.*, 2017), the protein with the highest similarity to the I-TASSER model. The putative AsCYP72A475 substrate recognition sites (SRS) were annotated based on CYP72A SRS as defined in a previous study (Prall *et al.*, 2016). The multiple sequence alignment Muscle (Edgar, 2004), was used to align the oat AsCYP72A475 protein sequence with those of *Catharanthus roseus* CYP72A1 and *Arabidopsis thaliana* CYP72A13 and the residues which aligned to those previously annotated were assigned as SRS 1-6.

References

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32**(5): 1792-1797.

Hsu MH, Baer BR, Rettie AE, Johnson EF. 2017. The crystal structure of cytochrome P450 4B1 (CYP4B1) monooxygenase complexed with octane discloses several structural adaptations for ω-hydroxylation. *Journal of Biological Chemistry* **292**(13): 5610-5621.

Mugford ST, Qi X, Bakht S, Hill L, Wegel E, Hughes RK, Papadopoulou K, Melton R, Philo M, Sainsbury F, et al. 2009. A serine carboxypeptidase-like acyltransferase is required for synthesis of antimicrobial compounds and disease resistance in oats. *Plant Cell* **21**(8): 2473-2484.

Prall W, Hendy O, Thornton LE. 2016. Utility of a phylogenetic perspective in structural analysis of CYP72A enzymes from flowering plants. *PLoS ONE* Sep 26;**11**(9):e0163024. doi: 10.1371/journal.pone.0163024.

Reed J, Stephenson MJ, Miettinen K, Brouwer B, Leveau A, Brett P, Goss RJM, Goossens A, O'Connell MA, Osbourn A. 2017. A translational synthetic biology platform for rapid access to gram-scale quantities of novel drug-like molecules. *Metabolic Engineering* **42**: 185-193.

Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ. 2009. Jalview version 2 – a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**(9): 1189-1191.

Werck-Reichhart D, Feyereisen R. 2000. Cytochromes P450: a success story. *Genome Biology* **1**(6): Reviews3003.1-3003.9 Epub Dec 8 2000.

Zhang Y. 2008. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* Jan 23;**9**:40. doi: 10.1186/1471-2105-9-40.