Supporting Information

Genetic Characterization of the Monodictyphenone Gene Cluster in *Aspergillus nidulans*

Yi-Ming Chiang[1,2], Edyta Szewczyk[3,8], Ashley D. Davidson[3], Ruth Entwistle[4], Nancy P. Keller[5,6], Clay C. C. Wang[2,7] & Berl R. Oakley[3,4]

[1]Graduate Institute of Pharmaceutical Science, Chia Nan University of Pharmacy and Science, Tainan 71710, Taiwan, ROC.

[2]Department of Pharmacology and Pharmaceutical Sciences, University of Southern California, School of Pharmacy, 1985 Zonal Avenue, Los Angeles, California 90089, USA.

[3]Department of Molecular Genetics, Ohio State University, 484 West 12th Avenue, Columbus, Ohio 43210, USA.

[4]Department of Molecular Biosciences, University of Kansas, 1200 Sunnyside Ave., Lawrence, KS 66045, USA.

[5]Department of Medical Microbiology and Immunology, University of Wisconsin-Madison, Madison, WI 53706, USA.

[6]Department of Bacteriology, University of Wisconsin-Madison, Madison, WI 53706, USA.

[7]Department of Chemistry, University of Southern California, College of Letters, Arts, and Sciences, Los Angeles, California 90089, USA.

[8]Current address, Research Center for Infectious Diseases, Röntgenring 11, D-97070, Würzburg, Germany.

Correspondence should be addressed to C.W. (clayw@usc.edu) and B.O. (boakley@ku.edu)

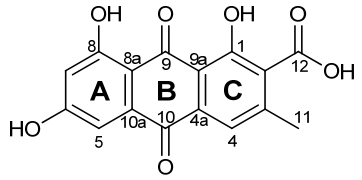
Table of Contents

Table S1. Primers used in this study.

primer	Sequence (5′→3′)
AN10039.4P1	GTA CAA CAC CGG CCT CTA GC
AN10039.4P2	ACC ACA CCC ATA CGC ATA CC
AN10039.4P3	CGA AGA GGG TGA AGA GCA TTG GAC ATG ACG ACA TGA TAC GG
AN10039.4P4	GCA TCA GTG CCT CCT CTC AGA CAG CTG CGT GAC CTT TCT TTT CC
AN10039.4P5	TGG CAG CAT CTA AGG ATT GG
AN10039.4P6	GAA GCC ATC CCC ACT AAT CC
AN10021.4P1	GTC ACC GAC CTG AAG TAC CC
AN10021.4P2	TGT CTT GTG AGT TGG GAT CG
AN10021.4P3	CGA AGA GGG TGA AGA GCA TTG CAA CCG ATA GAG CCT GAA CC
AN10021.4P4	GCA TCA GTG CCT CCT CTC AGA CAG GGA TAC AGT TCC GAA CAA GC
AN10021.4P5	TGA GGG ACT GAG GGT CTT CC
AN10021.4P6	GAC ACC ATG AGG GAC TGA GG
AN10049.4P1	ACC TCA ATT CCA ACG TCA GC
AN10049.4P2	AAA GTT GCC CTT GTG ACT GG
AN10049.4P3	CGA AGA GGG TGA AGA GCA TTG AGT GTC TAG GAC GGG AAG ACC
AN10049.4P4	GCA TCA GTG CCT CCT CTC AGA CAG TAG TTT CTG CGT CGG AAT CG
AN10049.4P5	CCA GCC TCG ACA ACA GAT CC
AN10049.4P6	GGC GCT GAC CTA TAA TTT GG
AN0146.4P1	CGC ACA GCT TCA TTC CTA CC
AN0146.4P2	CAA CAT GCC TCC AAT TAG CC
AN0146.4P3	CGA AGA GGG TGA AGA GCA TTG TGG AGA CAT TGG TGC TTT CC
AN0146.4P4	GCA TCA GTG CCT CCT CTC AGA CAG GTA AAA CCC GCC TTC ATA CG
AN0146.4P5	ATT CCG ACG CAG AAA CTA CC
AN0146.4P6	ATA TGC AGC CGA ACA TGA CC
AN0147.4P1	TGT AAC CAG TGT TGG GAC ACC
AN0147.4P2	GAA AGT GGC AGT GCA AGT CC
AN0147.4P3	CGA AGA GGG TGA AGA GCA TTG TTG GGT AGG GTC ATT GAA GC
AN0147.4P4	GCA TCA GTG CCT CCT CTC AGA CAG CAG TCG CAA TGT GAT TGA GC
AN0147.4P5	CAA TAC CTC AAC CAG GAG TCG
AN0147.4P6	CAG TGT TGG AGG ACA TGA GG
AN0148.4P1	CGA GGC AAC AGA CAA ATT CC
AN0148.4P2	TAT ACC ACC CCG AAC TCT GC
AN0148.4P3	CGA AGA GGG TGA AGA GCA TTG ATC AAT CGG GGG ATT ACA GC
AN0148.4P4	GCA TCA GTG CCT CCT CTC AGA CAG GAG TGG TCG GAG TCT TTT TCC
AN0148.4P5	ATG GAC CTT TGC GTG TTT CC
AN0148.4P6	GAG CAT GCG GTA GAA TTT CC
AN0149.4P1	GGT TCT GCG AGA TCT CAT CC
AN0149.4P2	GCG AGA TCT CAT CCA CTA ACG
AN0149.4P3	CGA AGA GGG TGA AGA GCA TTG AGT CTA GCC GAT GCT TTT GC
AN0149.4P4	GCA TCA GTG CCT CCT CTC AGA CAG ATT GGA TGG AGT GAG GTT GG
AN0149.4P5	CCC ATT CGA CCG ATA ACT CC
AN0149.4P6	ACG GAG GAG AAG GAC TTT GC
AN0150.4P1	TGG TGT GAA TTC AGC TTT CG
AN0150.4P2	ACA TAT GGT CAT GCG AGT GC
AN0150.4P3	CGA AGA GGG TGA AGA GCA TTG AAG TGA CAA GCG TCA GAT CG
AN0150.4P4	GCA TCA GTG CCT CCT CTC AGA CAG CCC ATC CTC ACT CAT CAA CC
AN0150.4P5	TTG ACT GAA CCC TGC TAG GC
AN0150.4P6	TAC TGG AAG CGC TGA TAT GC
AN10022.4P1	AAC AAC CTC GTG GAC TAC GC
AN10022.4P2	CTG TCC ACG GAG AAG AGT GG
AN10022.4P3	CGA AGA GGG TGA AGA GCA TTG TGG TTG ATG AGT GAG GAT GG
AN10022.4P4	GCA TCA GTG CCT CCT CTC AGA CAG TAG AGT CGC TTC GGG ACA TCA ACC
AN10022.4P5	TCC CAG CGA GCA GAA GAT AGA AG

AN10022.4P6	CTG GGA TTG GAG AAC GTA GC
AN10035.4P1	CAG CGA GAT CAA CCA TCA CC
AN10035.4P2	TTC TGC ATA TCA GCG CTT CC
AN10035.4P3	CGA AGA GGG TGA AGA GCA TTG GGG TTT CAG TGG AAC TGT CG
AN10035.4P4	GCA TCA GTG CCT CCT CTC AGA CAG AGA TGG ATT GTG TGC TGA GG
AN10035.4P5	GGA GTT CAT CGA GCG TAT CG
AN10035.4P6	CGG GTA CCG TAG CCT AAA CC
AN10038.4P1	TAC AAT CCC AGG CCA TTA GG
AN10038.4P2	GGA AGA AAT GCC TGA GCA AGC
AN10038.4P3	CGA AGA GGG TGA AGA GCA TTG CGA TAC GCT CGA TGA ACT CC
AN10038.4P4	GCA TCA GTG CCT CCT CTC AGA CAG TCG GTG GCG TTA AGA ATA GC
AN10038.4P5	GTA GTC ATG ACG GGG AAT GG
AN10038.4P6	CTC CAG ACA TGG AGG GAA GG
AN10044.4P1	TCC CGC AAC CTT CTT AAA CC
AN10044.4P2	CTC AAG GAC CCC ATC ATA CC
AN10044.4P3	CGA AGA GGG TGA AGA GCA TTG CGG GTA CCG TAG CCT AAA CC
AN10044.4P4	GCA TCA GTG CCT CCT CTC AGA CAG AGC CCT GAT CGA GGT TAA GG
AN10044.4P5	CAT CTC GGC AGT CTT TCT CG
AN10044.4P6	GCA CAG AGG TTT AGC ATC TCG
AN10023.4P1	TGA TCC AGA ATC TGC TCT CG
AN10023.4P2	CGC CTA CTG TCG AAA CAA GC
AN10023.4P3	CGA AGA GGG TGA AGA GCA TTG GGT AGA TGG TTG GGT TTT GC
AN10023.4P4	GCA TCA GTG CCT CCT CTC AGA CAG GGG TCT TGG CCA TCT AGT ACG
AN10023.4P5	CTC GGT CTG ACC ATT CTT GC
AN10023.4P6	GTG TTT TGC TCT TGC ACA GG
AN0153.4P1	GAG AAA GAC TGC CGA GAT GC
AN0153.4P2	GCC GAG ATG CTA AAC CTC TG
AN0153.4P3	CGA AGA GGG TGA AGA GCA TTG ATG ATG CTT CCA GGA TCA GC
AN0153.4P4	GCA TCA GTG CCT CCT CTC AGA CAG CCG TCA GTC AGT CAA AGT GG
AN0153.4P5	CTG CCT CCT TTA CCC GTC TCC
AN0153.4P6	AGC CTT GCT GCC TCC TTT ACC
alcA_AN10021.4P1	GTC ACC GAC CTG AAG TAC CC
alcA_AN10021.4P2	TGT CTT GTG AGT TGG GAT CG
alcA_AN10021.4P3	CGA AGA GGG TGA AGA GCA TTG CAA CCG ATA GAG CCT GAA CC
alcA_AN10021.4P4	ATC CTA TCA CCT CGC CTC AAA ATG ATG TCT AGT CTA TCC GAC C
alcA_AN10021.4P5	GAA GGT CGT CGT GTT TGT GG
alcA_AN10021.4P6	GGT GTG GTT GTG GCT AGA GG

Blue and red sequences are tails that anneal to the *A. fumigatus pyroA* (*AfpyroA*) fragment during fusion PCR. Green sequence is a tail that anneals to the *alcA* promoter fragment during fusion PCR.



9

Table S2. NMR data for endocrocin **9** (400 and 100 MHz in DMSO-d_6 and CD_3OD).

	DMSO-d_6[a]	DMSO-d_6[b]	DMSO-d_6[c]	DMSO-d_6		CD_3OD	
position	δ_C	δ_C	δ_C	δ_C	δ_H	δ_C	δ_H
1	157.9	157.9	157.8	no[e]	—	159.9	—
2	130.3	129.1	130.5	129.1	—	132.2	—
3	143.6	144.0	143.6	145.4	—	145.6	—
4	120.4	120.5	120.5	120.0	7.41 (1H, s)	122.1	7.58 (1H, s)
4a	134.7[d]	130.5	135.1	135.0[d]	—	136.9[d]	—
5	108.9	108.9	108.9	108.2	7.09 (1H, d, J = 2.4 Hz)	110.4	7.15 (1H, d, J = 2.4 Hz)
6	165.7	165.7	165.7	165.0	—	167.6	—
7	107.8	107.9	108.0	108.2	6.59 (1H, d, J = 2.4 Hz)	109.2	6.54 (1H, d, J = 2.4 Hz)
8	164.5	164.5	164.5	164.3	—	166.8	—
8a	108.8	109.1	109.1	109.7	—	110.7	—
9	189.1	no[e]	167.0[f]	188.7	—	191.8	—
9a	113.5	114.9	113.9	114.7	—	115.4	—
10	180.5	181.8	181.1	181.9	—	182.7	—
10a	132.2[d]	no[e]	132.5	133.1[d]	—	134.5[d]	—
11	19.7	19.4	19.7	21.2	2.48 (3H, s)	20.4	—
12	167.0	no[e]	189.4[f]	168.1	—	170.0	—
OH	—	—	—	—	12.59 (1H, s)	—	—

[a]Data obtained from Kurobane et al., *J. Antibiot. (Tokyo)* **32**, 1256-1266 (1979). [b]Data obtained from Waser et al., *Tetrahedron Lett.* **46**, 2377-2380 (2005). [c]Data obtained from Awakawa et al., *Chem. Biol.* **16**, 613-623 (2009). [d]Values in the same column may be interchanged. [e]Not observed. [f]These two signals might be misassigned. Some differences in the δ_C with reported literature values might be due to the C-1 hydroxy-assisted tautomerization of the carboxylic acid.

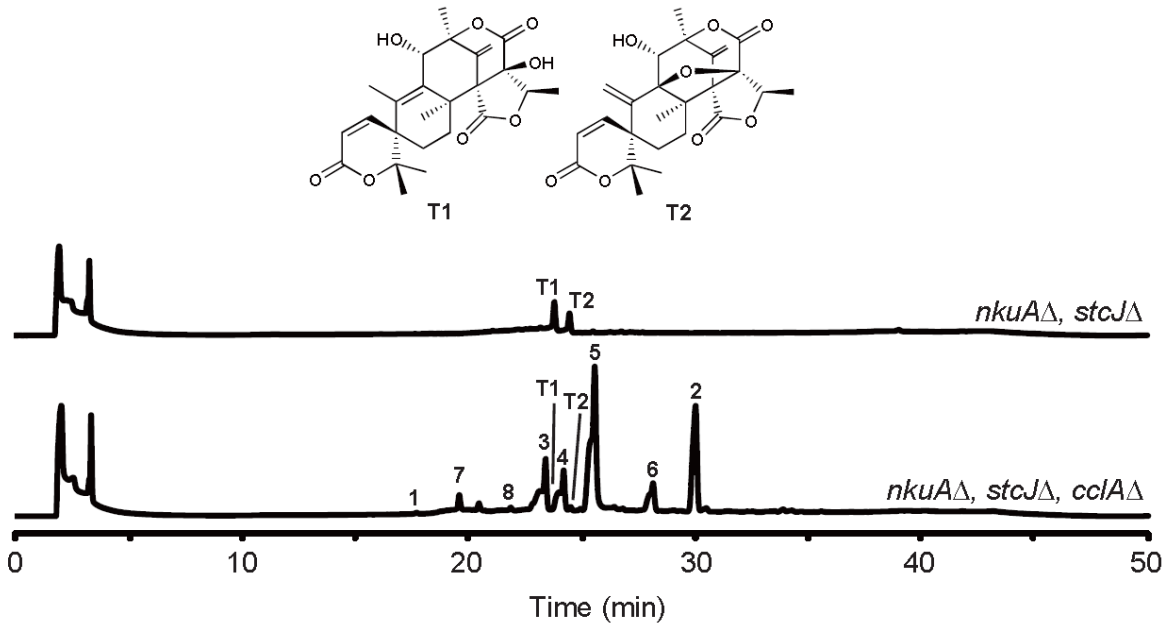
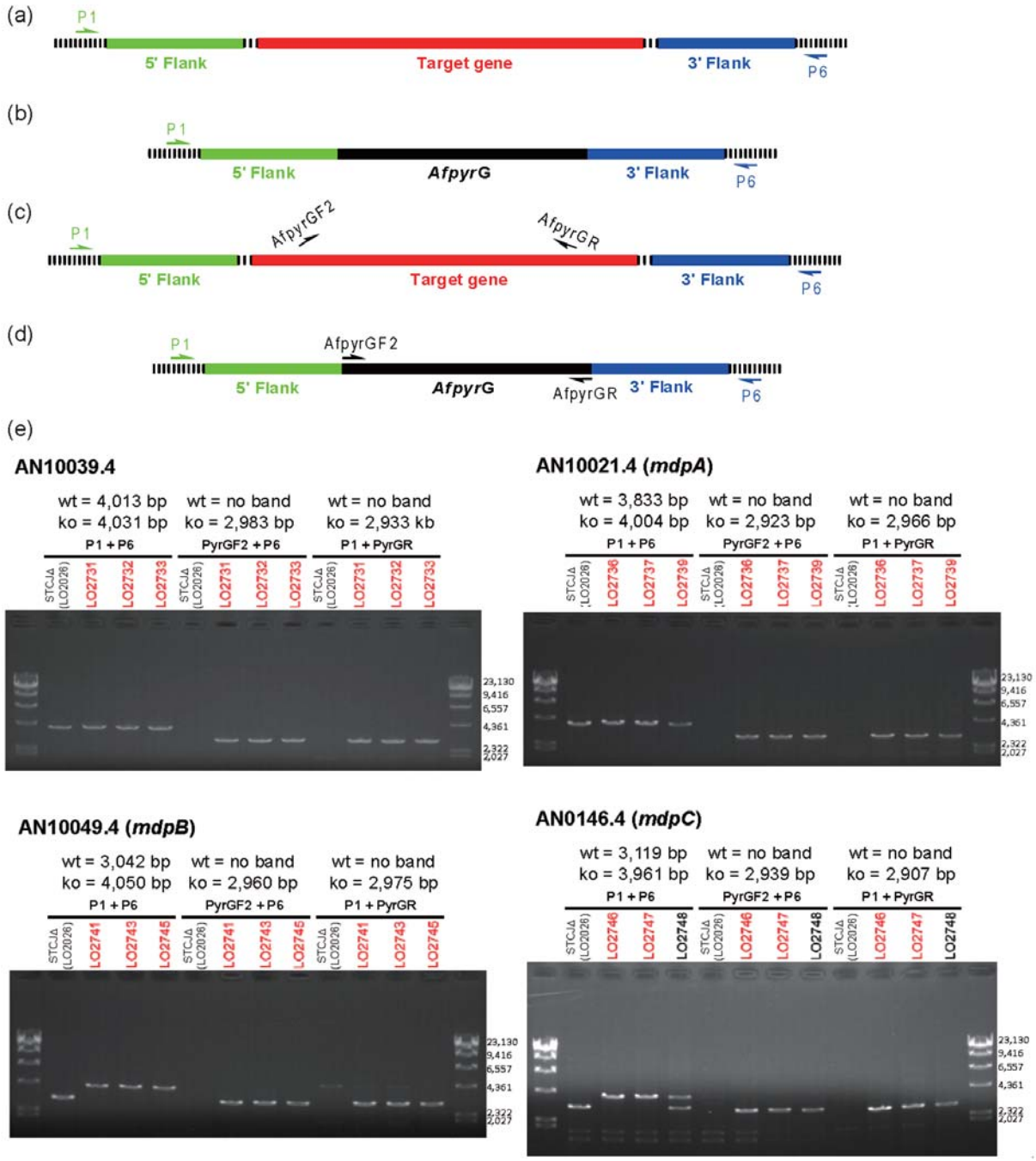
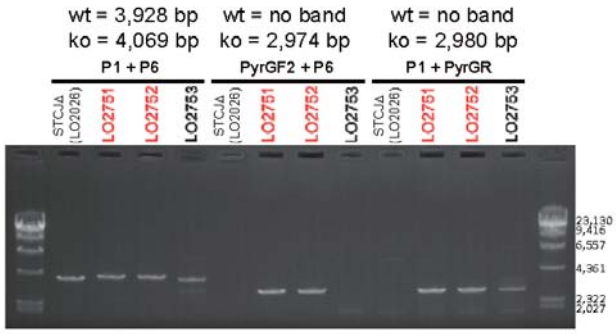


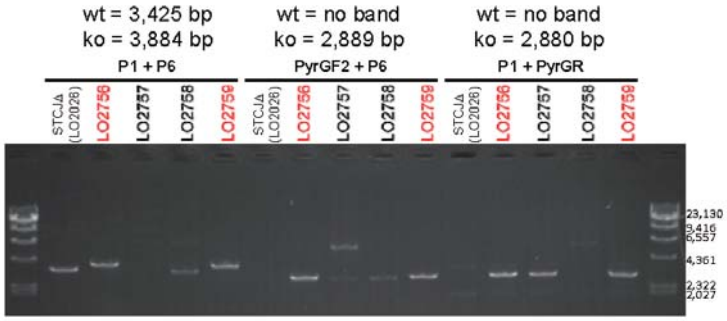
Figure S1. Total UV scan HPLC profiles of *A. nidulans nkuA*Δ, *stcJ*Δ double knockout and *nkuA*Δ, *stcJ*Δ, *cclA*Δ triple knockout strains. The *nkuA*Δ, *stcJ*Δ double knockout strain produces mainly meroterpenoids (T1 and T2). The *nkuA*Δ, *stcJ*Δ, *cclA*Δ triple knockout strain produces meroterpenoids as well as monodictyphenone (**1**), emodin (**2**), and emodin derivatives (**3** – **8**). The retention times of some compounds are shifted slightly when compared with **Figure 3**. This was due to the use of different batches of HPLC columns.


(a)
P1
5' Flank
Target gene
3' Flank
P6
(b)
P1
5' Flank
AfpyrG
3' Flank
P6
(c)
P1
AfpyrGF2
AfpyrGR
5' Flank
Target gene
3' Flank
P6
(d)
P1
AfpyrGF2
5' Flank
AfpyrG
AfpyrGR
3' Flank
P6
(e)
AN10039.4
wt = 4,013 bp
ko = 4,031 bp
wt = no band
ko = 2,983 bp
wt = no band
ko = 2,933 kb
P1 + P6
PyrGF2 + P6
P1 + PyrGR
STC⊿A (LO2026)
LO2731
LO2732
LO2733
23,130
9,416
6,557
4,361
2,322
2,027
AN10021.4 (mdpA)
wt = 3,833 bp
ko = 4,004 bp
wt = no band
ko = 2,923 bp
wt = no band
ko = 2,966 bp
P1 + P6
PyrGF2 + P6
P1 + PyrGR
STC⊿A (LO2026)
LO2736
LO2737
LO2739
23,130
9,416
6,557
4,361
2,322
2,027
AN10049.4 (mdpB)
wt = 3,042 bp
ko = 4,050 bp
wt = no band
ko = 2,960 bp
wt = no band
ko = 2,975 bp
P1 + P6
PyrGF2 + P6
P1 + PyrGR
STC⊿A (LO2026)
LO2741
LO2743
LO2745
23,130
9,416
6,557
4,361
2,322
2,027
AN0146.4 (mdpC)
wt = 3,119 bp
ko = 3,961 bp
wt = no band
ko = 2,939 bp
wt = no band
ko = 2,907 bp
P1 + P6
PyrGF2 + P6
P1 + PyrGR
STC⊿A (LO2026)
LO2746
LO2747
LO2748
23,130
9,416
6,557
4,361
2,322
2,027

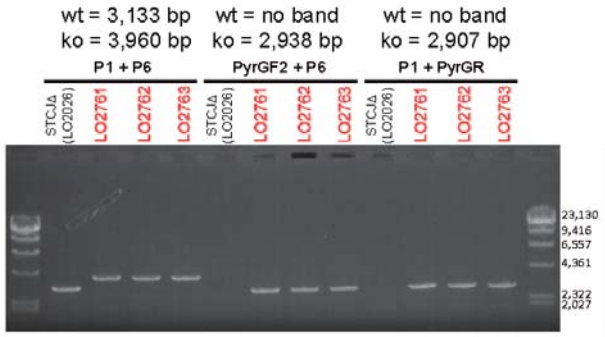
AN0147.4 (*mdpD*)



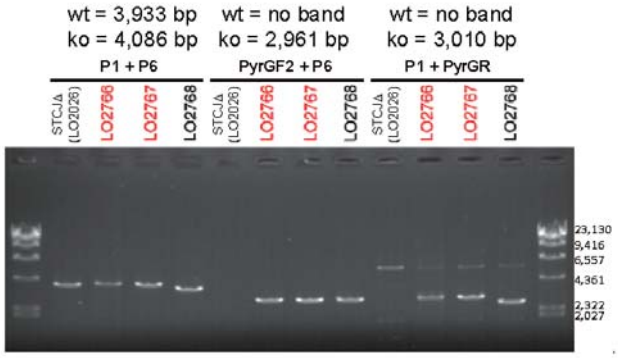
AN0148.4 (*mdpE*)



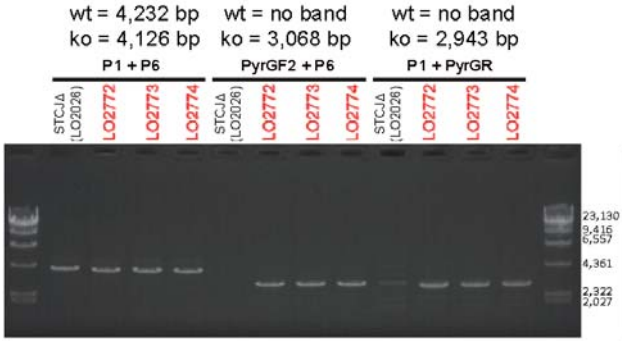
AN0149.4 (*mdpF*)



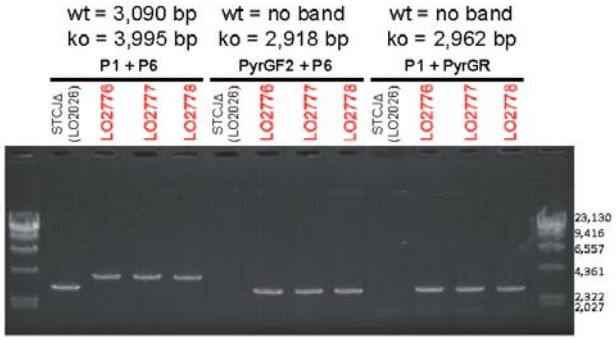
AN10022.4 (*mdpH*)



AN10035.4 (*mdpI*)



AN10038.4 (*mdpJ*)



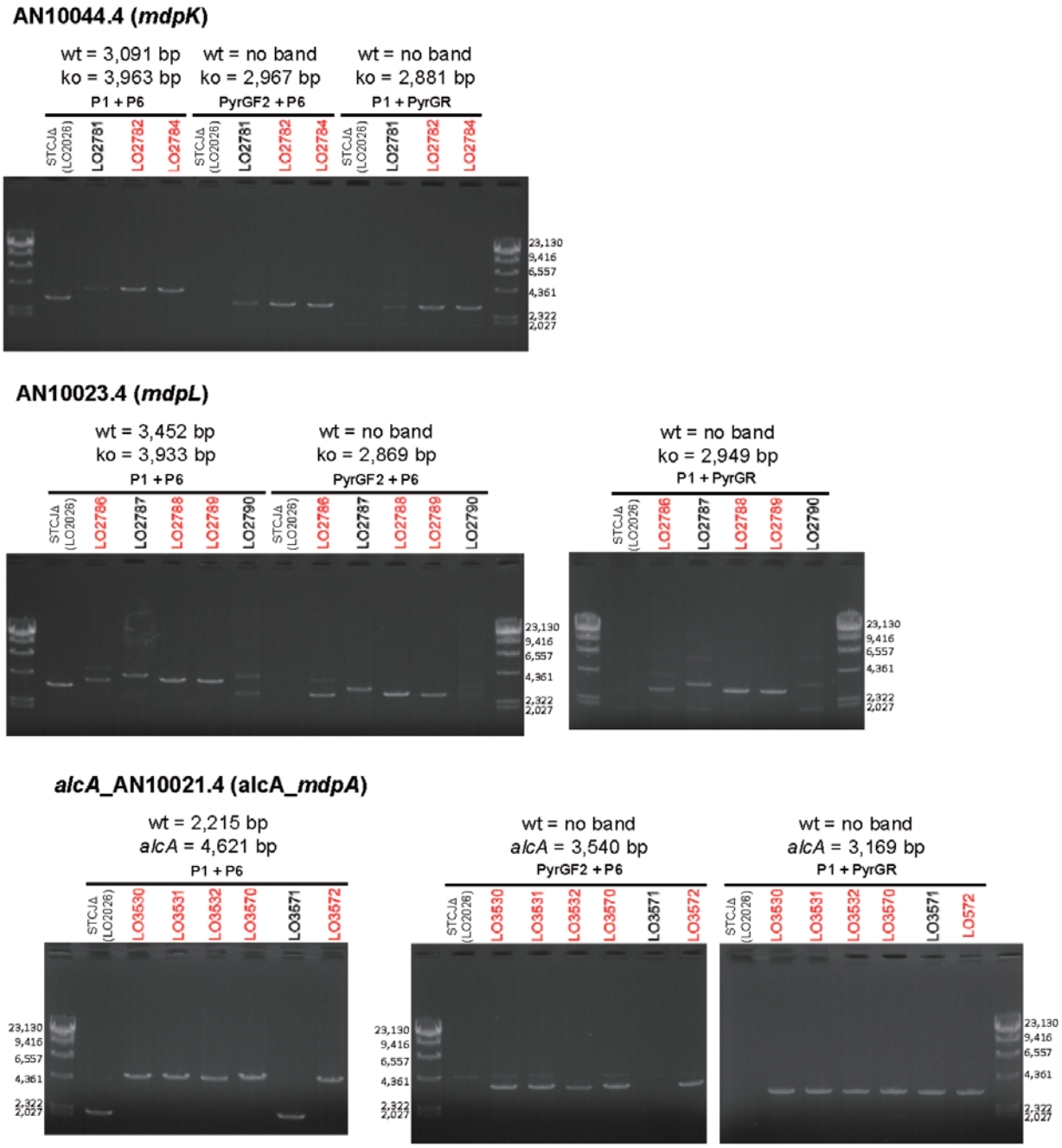


Figure S2. Diagnostic PCR strategy and results. We used two redundant strategies to determine if the target gene had been deleted by replacement with *AfpyrG*. In one strategy, DNA from transformants is amplified with two primers, P1 from the chromosomal region just outside of the 5' flank of the transforming DNA fragment and P6 from just outside of the 3' flank. If the target gene is different in size from the *AfpyrG* gene which was used as a selectable

marker for transformation, the PCR fragment amplified from a correct transformant (b) will be different in size from the fragment amplified if the target gene is intact (a). In some instances the target gene and the *AfpyrG* cassette will be almost the same size and this strategy will not produce definitive results. In the second strategy, P1 and P6 are used with primers specific to the *AfpyrG* cassettes. For example, if the target gene has been replaced by the *AfpyrG* gene (d), P1 and AfpyrGR will amplify a fragment of a predictable size. If the target gene has not been replaced (c), the AfpyrGR primer will not anneal and there will be no specific amplification. Likewise AfpyrGF2 and P6 are used in combination and amplification will only occur if the target gene has been replaced by *AfpyrG*. For promoter replacements of *mdpA*, *AfpyroA* was used as a selectable marker. For ease of creating fusion PCR constructs, the *AfpyroA* cassette had been constructed such that it contained AfpyrGF2 and AfpyrGR sequences that allowed verification with the same strategies as for constructs using *AfpyrG* as a selectable marker. We additionally verified the *mdpA* promoter replacements with *AfpyroA*-specific primers (data not shown). (e) Results of diagnostic PCR for mdp cluster deletions. Strain numbers labeled in red are correct transformants.

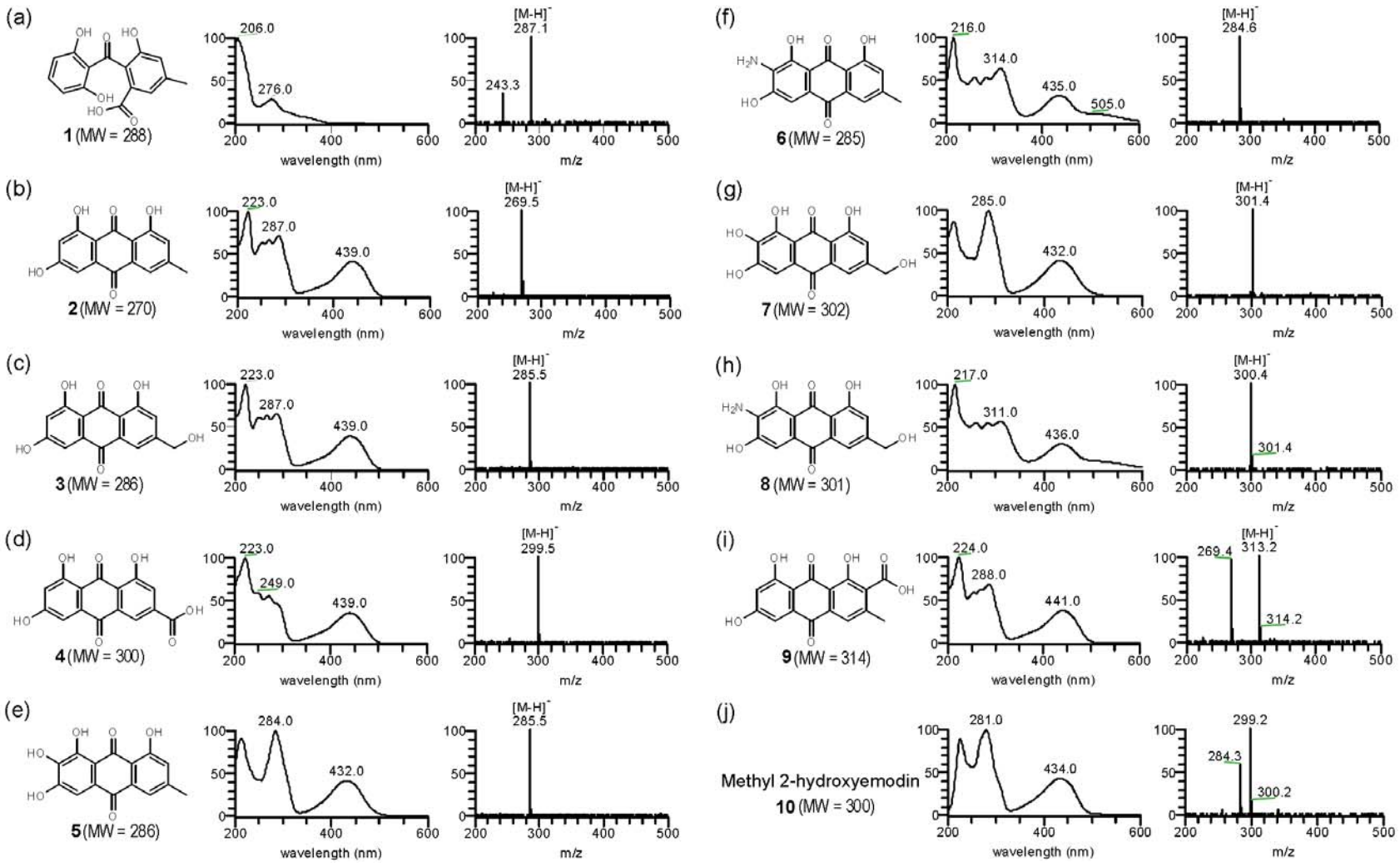


Figure S3. UV-Vis and ESIMS spectra in negative mode of monodictyphenone (**1**), emodin (**2**), and emodin derivatives (**3** – **10**).

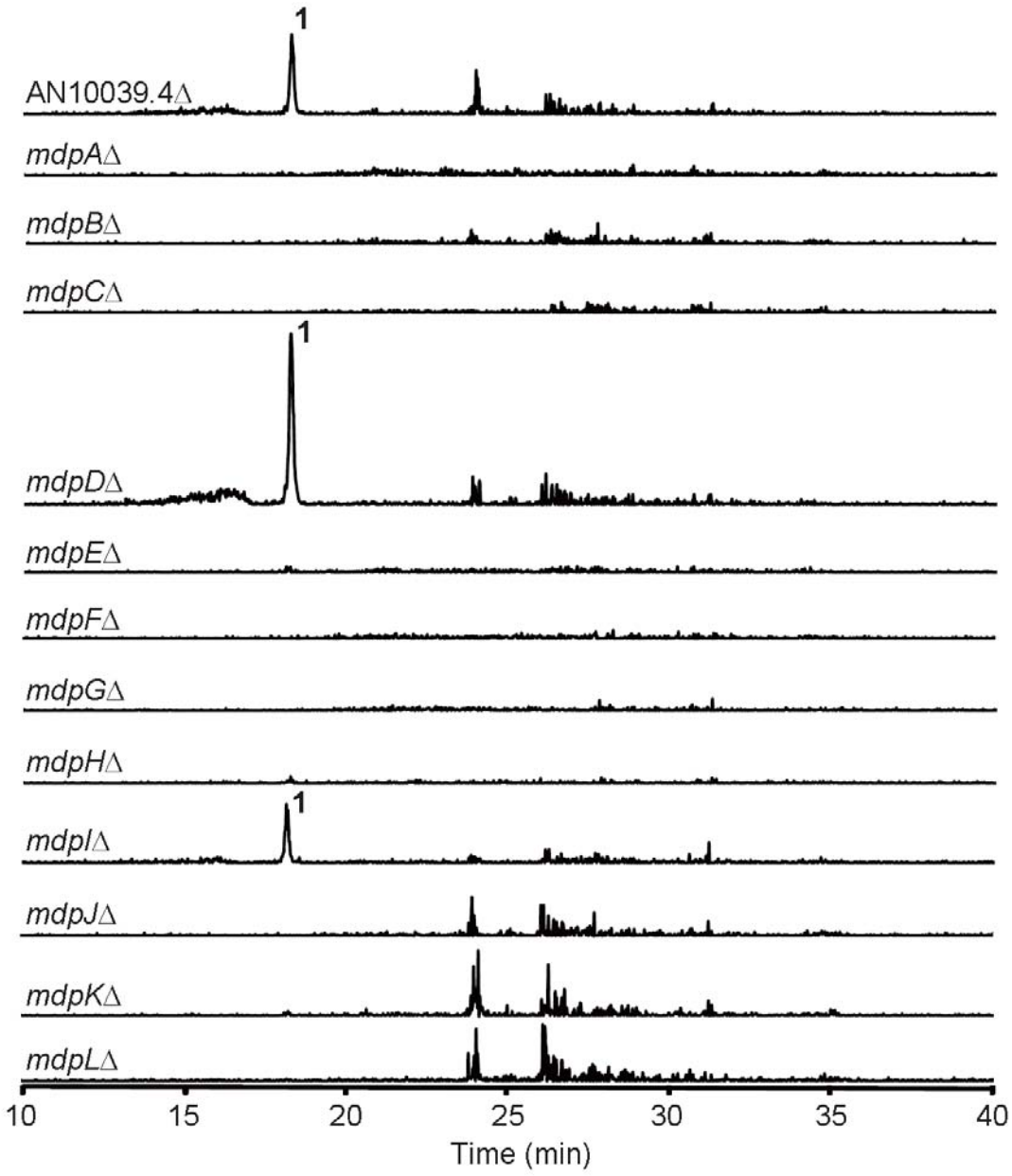


Figure S4. Analysis of the effects on monodictyphenone (**1**) production of gene deletions in the mdp cluster by negative mode extracted ion chromatograms (EIC) at *m*/*z* 287. All strains have *nku*Δ, *stcJ*Δ, and *cclA*Δ background as in Figure 3. The Y axis of each profile was at the same order of magnitude.

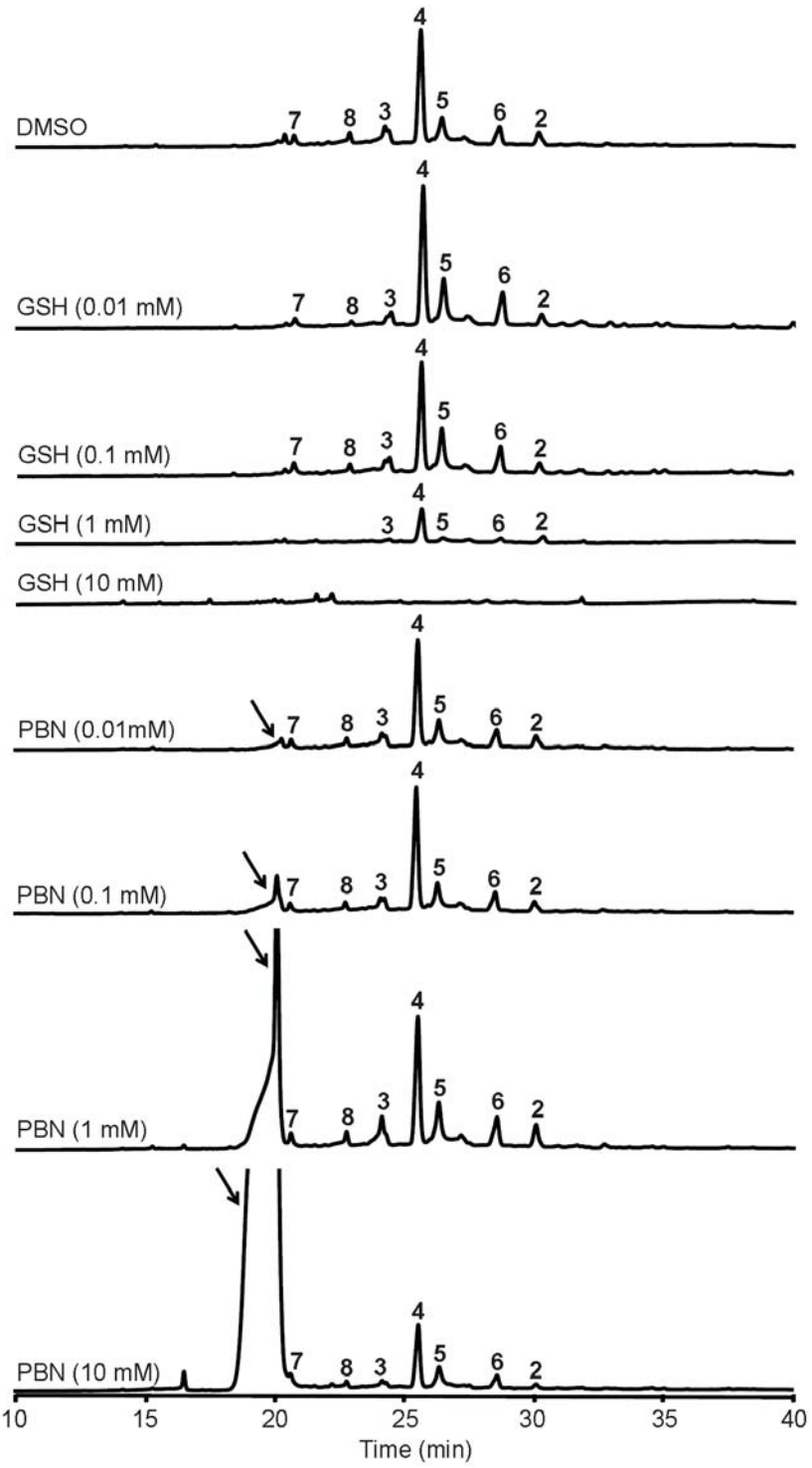


Figure S5. Metabolite profiles of *A. nidulans* strain LO2051 (*nkuA*Δ, *stcJ*Δ, *cclA*Δ) incubated with vehicle control (0.5% DMSO), and various concentration of glutathione (GSH) and *N*-*tert*-butyl-α-phenylnitrone (PBN) at UV 276 nm. The arrow indicated the location of PBN. The Y axis of each profile was at the same order of magnitude.

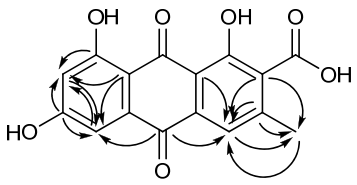


Figure S6. HMBC correlations of endocrocin (**9**).

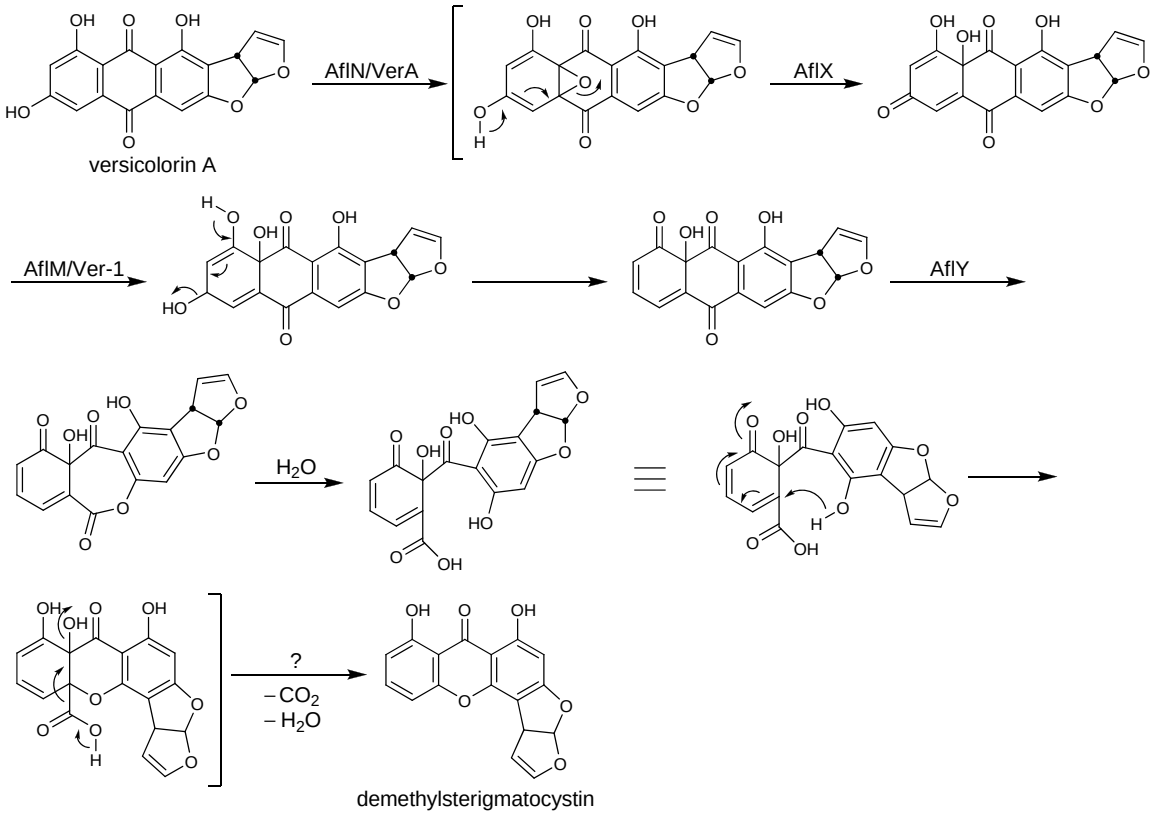


Figure S7. Possible pathway for conversion of versicolorin A to demethylsterigmatocystin.

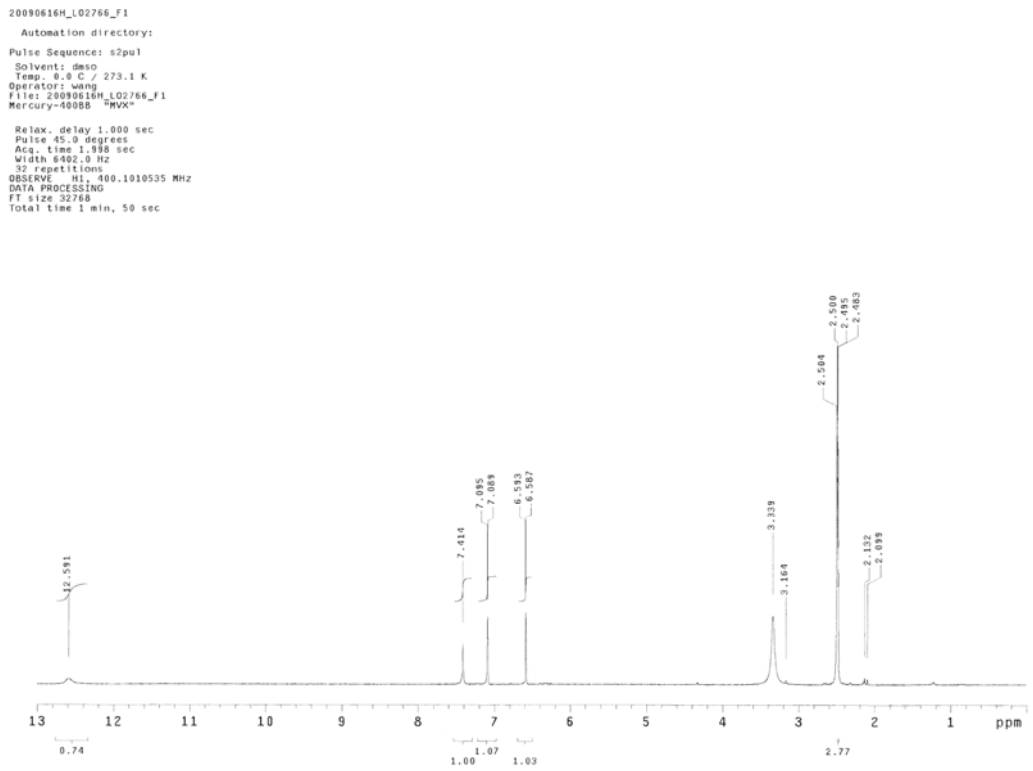


Figure S8. ^{1}H NMR spectrum of endocrocin (**9**) in DMSO-d_6.

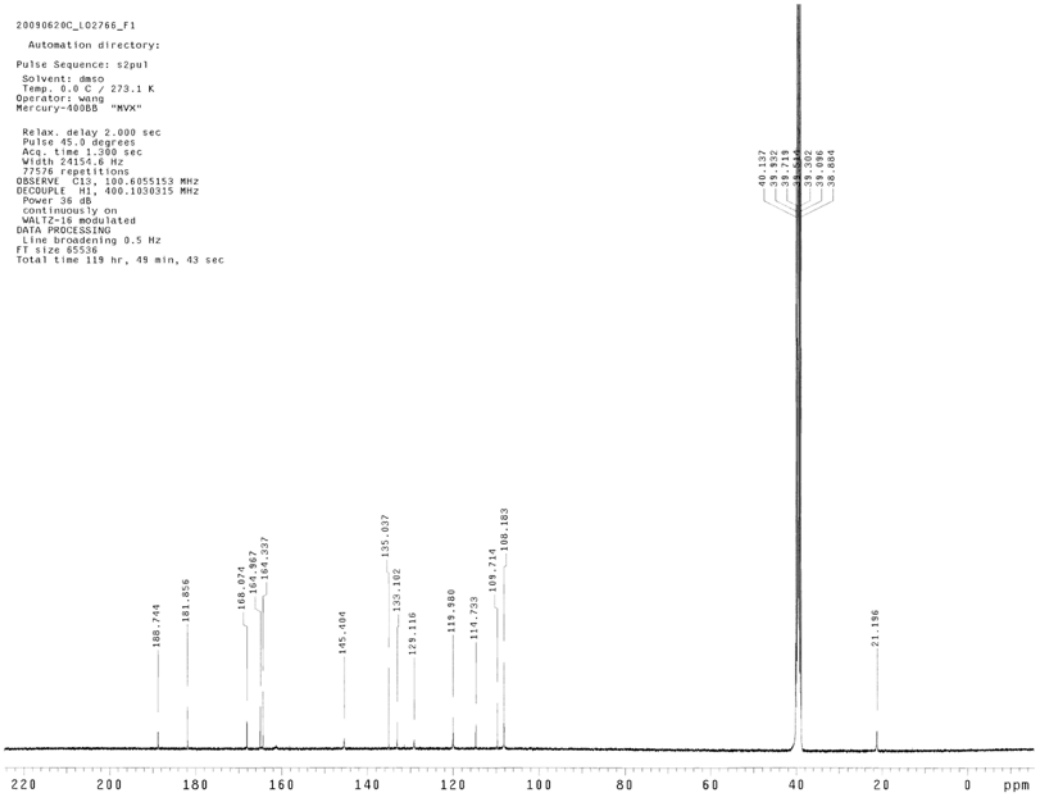


Figure S9. ^{13}C NMR spectrum of endocrocin (**9**) in DMSO-d_6.

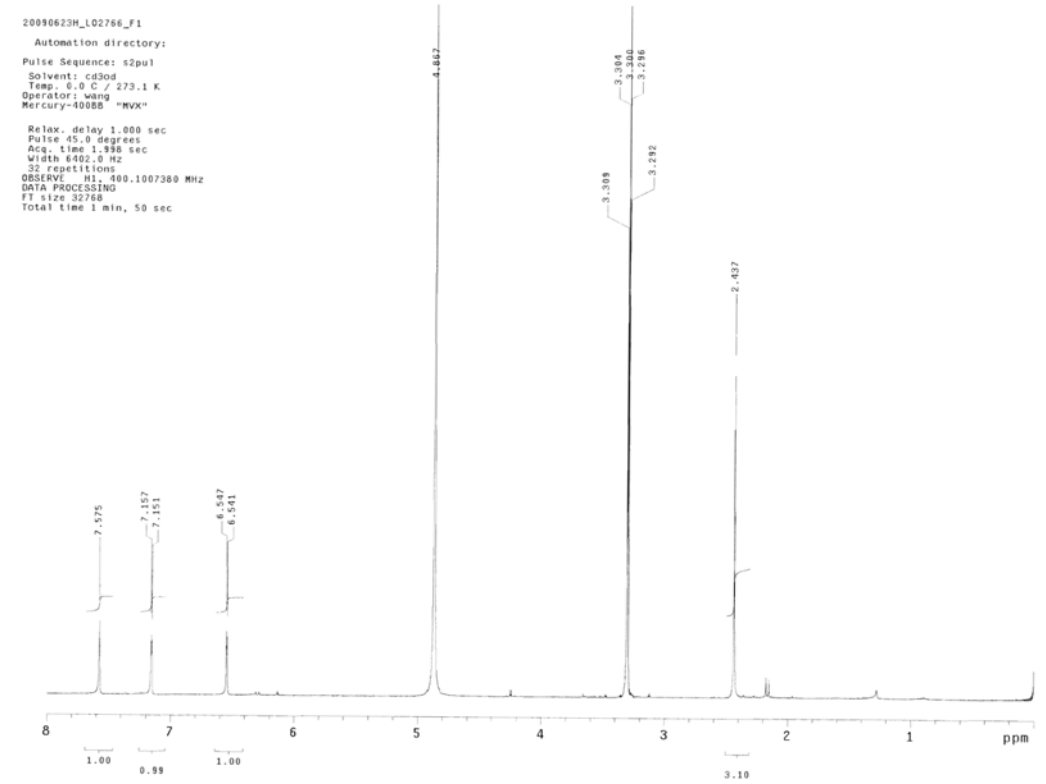


Figure S10. ^{1}H NMR spectrum of endocrocin (**9**) in CD_3OD.

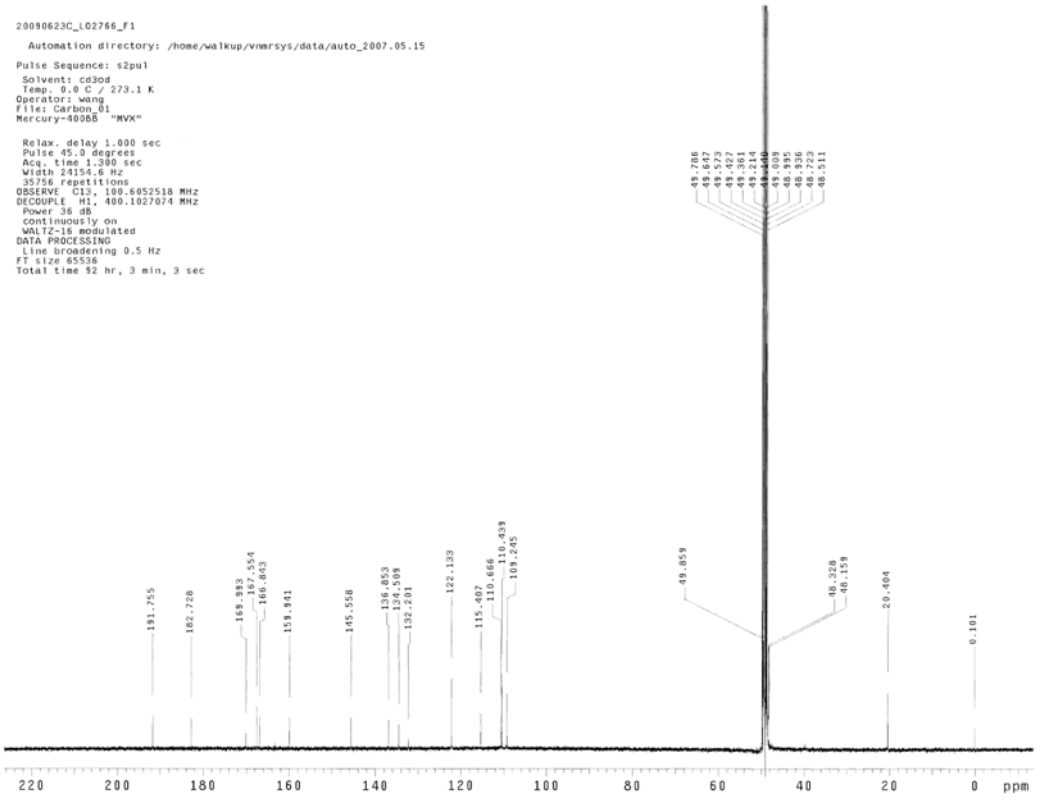


Figure S11. ^{13}C NMR spectrum of endocrocin (**9**) in CD_3OD.