JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Article

pubs.acs.org/JACS

Illuminating the Diversity of Aromatic Polyketide Synthases in *Aspergillus nidulans*

Manmeet Ahuja,[†] Yi-Ming Chiang,*[,‡,⊥] Shu-Lin Chang,[§,⊥] Mike B. Praseuth,[⊥] Ruth Entwistle,[†] James F. Sanchez,[⊥] Hsien-Chun Lo,[⊥] Hsu-Hua Yeh,[⊥,∥] Berl R. Oakley,*[,†] and Clay C. C. Wang*[,⊥,#]

[†]Department of Molecular Biosciences, University of Kansas, 1200 Sunnyside Avenue, Lawrence, Kansas 66045, United States

[‡]Graduate Institute of Pharmaceutical Science, Chia Nan University of Pharmacy and Science, Tainan 71710, Taiwan, Republic of China

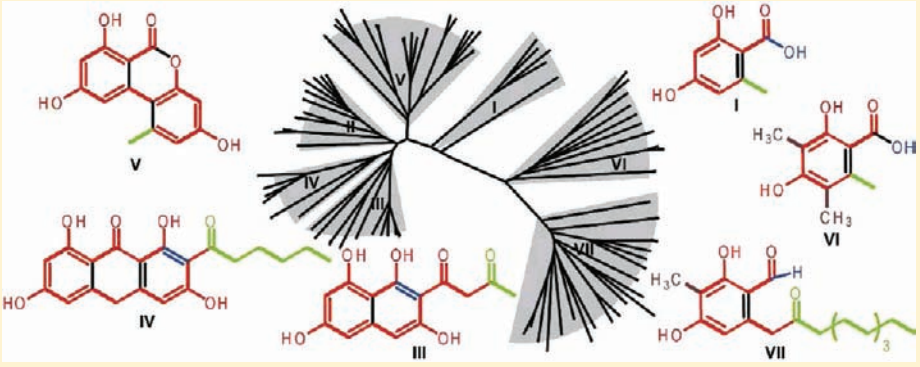
[§]Department of Biotechnology, Chia Nan University of Pharmacy and Science, Tainan 71710, Taiwan, Republic of China

[⊥]Department of Pharmacology and Pharmaceutical Sciences, School of Pharmacy, University of Southern California, 1985 Zonal Avenue, Los Angeles, California 90089, United States

[∥]Department of Biological Science and Technology, National Chiao Tung University, Hsinchu 30010, Taiwan, Republic of China

[#]Department of Chemistry, College of Letters, Arts, and Sciences, University of Southern California, Los Angeles, California 90089, United States

Supporting Information



ABSTRACT: Genome sequencing has revealed that fungi have the ability to synthesize many more natural products (NPs) than are currently known, but methods for obtaining suitable expression of NPs have been inadequate. We have developed a successful strategy that bypasses normal regulatory mechanisms. By efficient gene targeting, we have replaced, *en masse*, the promoters of nonreducing polyketide synthase (NR-PKS) genes, key genes in NP biosynthetic pathways, and other genes necessary for NR-PKS product formation or release. This has allowed us to determine the products of eight NR-PKSs of *Aspergillus nidulans*, including seven novel compounds, as well as the NR-PKS genes required for the synthesis of the toxins alternariol (**8**) and cichorine (**19**).

INTRODUCTION

Fungal natural products (NPs) are an important source of medically important compounds, and sequencing projects have revealed that many fungal genomes contain large numbers of clusters of genes that are predicted to encode NP biosynthetic pathways.[1,2] The number of predicted NP biosynthetic pathways for each species far exceeds the known number of natural products, and there is great variation in NP gene clusters among species. Fungi, thus, have the capacity to produce many more NPs than are currently known and precedence indicates that many of them have remarkably useful medical activities. Most fungal biosynthetic pathways are cryptic, however, producing no products under normal laboratory growth conditions.[3,4]

Several approaches have been employed to obtain expression of orphan pathways.[3−5] One approach has been to alter genes involved in chromatin packing to induce expression.[6] While this has worked to a limited extent, this approach, so far has not lead to the expression of most NP gene clusters. Other approaches such as altering media[7] or coculturing with bacteria[8] have produced some success, but the great majority of NP gene clusters have not responded to these approaches. Clearly, to unlock the treasure house of natural products that fungi can produce, we need to have a more generally successful approach to fungal NP production.

One promising approach takes advantage of the fact that fungal natural product gene clusters often contain transcription factors that govern expression of all genes in the cluster.[9]

Received: February 18, 2012
Published: April 17, 2012

   dx.doi.org/10.1021/ja3016395 | *J. Am. Chem. Soc.* 2012, 134, 8212−8221

Table 1. Promoter-Replaced Putative Transcription Factors Associated with NP Gene Clusters[a]

PKS or NRPS	promoter-replaced putative transcription factors	product induced
ANID_00150.1 (NR-PKS, MdpG)	ANID_00148.1 (*MdpE*)	emodin derivatives[6,10]
	ANID_10021.1 (*MdpA*)	nd[6,10]
	ANID_00148.1 (*MdpE*) and ANID_10021.1 (*MdpA*)	emodin derivatives[6,10]
ANID_00523.1 (NR-PKS)	AN0533.3	nd
ANID_01034.1 (NR-PKS, AfoE), ANID_01036.1 (HR-PKS, AfoG)	ANID_01029.1 (*AfoA*)	asperfuranone[11]
ANID_01680.1 (NRPS-like)	ANID_01678.1	nd
ANID_02032.1 (NR-PKS), ANID_02035.1 (HR-PKS)	ANID_02025.1	low
	ANID_02026.1	low
	AN2036.3	low
	ANID_02025.1 and AN2036.3	low
ANID_02924.1 (NRPS-like)	ANID_02919.1	nd
ANID_03386.1 (NR-PKS), ANID_03396.1 (NRPS-like)	AN3385.3	low
	ANID_03391.1	nd
	AN3385.3 and ANID_03391.1	low
ANID_03495.1 (NRPS), ANID_03496.1 (NRPS)	AN3501.3	nd
	ANID_03502.1	nd
	AN3501.3 and ANID_03502.1	low
	ANID_03506.1	nd
ANID_10486.1 (NRPS)	AN3911.3	nd
	ANID_10491.1	nd
ANID_06431.1 (HR-PKS)	AN6430.3	nd
ANID_06791.1 (HR-PKS)	ANID_06788.1	nd
	ANID_06790.1	nd
ANID_07071.1 (NR-PKS)	ANID_07061.1	nd
	AN7073.3	nd
	ANID_07061.1 and AN7073.3	nd
ANID_08105.1 (NRPS-like)	AN8103.3	nd
	ANID_08111.1	nd
	AN8103.3 and ANID_08111.1	nd
ANID_08412.1 (PKS-NRPS, ApdA)	ANID_08414.1 (*ApdR*)	aspyridones[9]
ANID_08513.1 (NRPS, TdiA)	ANID_08506.1	low
	ANID_08509.1	nd
ANID_08910.1 (HR-PKS)	AN8916.3	nd
	AN8918.3	nd
	AN8916.3 and AN8918.3	nd
ANID_11191.1 (HR-PKS), ANID_09226.1 (NRPS)	ANID_09221.1	nd
	ANID_09236.1	low
	ANID_09221.1 and ANID_09236.1	low

[a]In some cases, there are two or more putative transcription factors associated with a single NP gene cluster. In those cases, we replaced the promoters of each transcription factor individually with the *alcA* promoter and both of the transcription factors in combination. Named products are from previous studies in our laboratories and others. The remainder of the promoter replacements were carried out in this study (Supplementary Table S3 and Figure S1, Supporting Information). "nd" indicates no product was detected. "low" indicates that detectable products were produced but in amounts too low to allow analysis. The annotations of some transcription factors changed over the course of these experiments. Genes beginning with ANID are based on the Broad Institute annotation (http://www.broadinstitute.org/annotation/genome/aspergillus_group/MultiHome.html). Genes beginning with AN are based on the .3 annotation on the CADRE database (http://www.cadre-genomes.org.uk/index.html).

Replacing the promoters of these transcription factors with regulatable promoters has allowed the induction of expression of three clusters in *Aspergillus nidulans* that encode biosynthetic pathways for aspyridones,[9] emodin derivatives,[6,10] and asperfuranone[11] (Table 1). Since this approach has been successful for us and for others, we wished to determine whether this approach was generally applicable. We found, however, that in most cases, induction of these transcription factors did not result in induction of useful amounts of NPs. We consequently developed an alternative strategy, completely bypassing normal regulation, in which we directly replaced the promoters of nonreducing polyketide synthase (NR-PKS) genes, key genes of NP biosynthetic pathways, as well as other genes required for compound production or release.[12,13] Because we did not replace the promoters of genes that may modify the products of the NR-PKS, we did not anticipate that we would obtain production of the final products of the NP biosynthetic pathways. We did, however, hope to identify novel compounds and develop a better understanding of NR-PKS function.

We report here that this approach is highly successful. It has allowed us to complete the determination of the products of eight NR-PKS genes of *A. nidulans*. We have discovered seven novel compounds, and our findings have allowed us to determine the NR-PKS genes required for synthesis of cichorine (**19**), a phytotoxin produced by *A. nidulans*,[14] and of alternariol (**8**),[15] an important toxic contaminant of cereals and fruits not previously known to be produced by *Aspergillus sp*. Our approach should allow the discovery of products of other NP gene clusters and the production of large amounts of NPs from target clusters, and facilitate the prediction of chemical structures of the products of novel NR-PKS identified in fungal genome projects.

RESULTS

We took advantage of progress in generating transforming fragments by fusion PCR and in gene targeting in *A. nidulans*[16,17] to replace the promoters of all putative transcription factors (TFs) associated with NP clusters in *A. nidulans* that have not previously been studied (Table 1). We replaced them with the regulatable promoter from the *alcA* gene[18] using the strategy shown in Figure 1A. The *alcA* promoter allows us to repress expression of genes while the strain germinates and grows, and then induce the expression to very high levels. This allows us, in principle, to induce production of compounds that might be toxic to the producing strain. In instances in which the TFs are predicted on the basis of sequence to involve proteins encoded by two genes, we replaced the promoters of both genes. In some cases, multiple putative TFs are associated with a single NP cluster, and in those cases, we replaced the promoters of each TF. In all, we created 33 new combinations of single and multiple TF promoter replacements for NP clusters for which the products are unknown. We then induced expression of the

dx.doi.org/10.1021/ja3016395 | J. Am. Chem. Soc. 2012, 134, 8212−8221

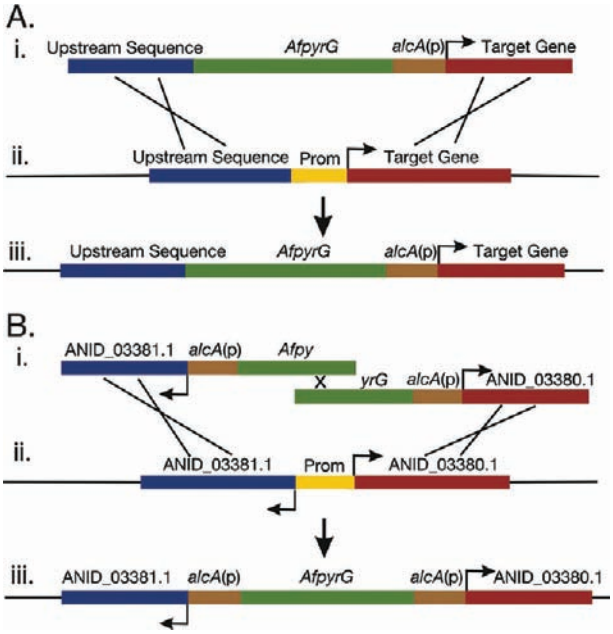


Figure 1. Promoter replacement strategies. (A) Single promoter replacement strategy: (i) A transforming sequence is generated by fusion PCR. It consists of a sequence upstream of the target gene, a selectable marker such as the *pyrG* gene from *Aspergillus fumigatus* (*AfpyrG*), the *alcA* promoter (*alcA*(p)), and all or a portion of the target gene. In the present study, the upstream sequence and the portion of the target gene were about 1 kb each in length. Transformation with this fragment results in homologous recombination with the chromosomal locus (ii) and replacement of the promoter of the target gene with a selectable marker and the *alcA* promoter, placing the target gene under the control of the *alcA* promoter (iii). This strategy was used to replace the promoters of transcription factors, NR-PKSs, and other genes of the NP gene clusters. (B) Dual promoter replacement strategy: Fatty acid synthase subunits such as ANID_03381.1 and ANID_03380.1 are generally divergently transcribed in *A. nidulans*. Two fragments were created by fusion PCR (i). One contained a portion of one target gene (ANID_03381.1) fused to *alcA*(p) and a portion of the *AfpyrG* gene. The second contained a second, overlapping, portion of the *AfpyrG* gene and *alcA*(p) fused to the second target gene (ANID_03380.1). Upon transformation, recombination between the target genes on the transforming fragments and the chromosomal genes (ii) and between the overlapping fragments of *AfpyrG* results in both target genes being driven by *alcA*(p) (iii).

TFs and analyzed organic extracts for production of new NPs by HPLC-DAD-MS (Supplementary Figure S1, Supporting Information). Among 17 clusters studied, induction of TFs in 9 did not result in production of any detectable new NPs; and in 5 clusters, induction of one or more TFs resulted in the production of new NPs but in amounts too small to be isolated and structurally characterized by NMR (Table 1 and Supplementary Figure S1, Supporting Information). This strategy, thus, does not result in adequate up-regulation of most NP clusters that contain transcription factors.

Clearly, to unlock the treasure house of natural products that fungi can produce, we needed to have a more generally successful approach for fungal NP production. Given the efficiency of gene targeting that is now possible in *A. nidulans*, we reasoned that the systematic replacement of native promoters of unknown NP genes in *A. nidulans* with the regulatable *alcA* promoter would provide expression levels suitable to identify new NPs. To determine whether bypassing normal NP regulatory mechanisms was feasible and practical, we focused our attention on NR-PKSs.

Fungal polyketide synthases (PKSs)[12,19] are key NP genes, and they are abundant in the *A. nidulans* genome. We analyzed all 29 *A. nidulans* PKS genes and classified them into 14 NR-PKSs, 13 highly reducing PKSs (HR-PKSs), one HR-PKS lacking an enoyl reductase (ER) domain, and one hybrid PKS–nonribosomal peptide synthetase, based on phylogeny and their domain architectures (Supplementary Figure S2, Supporting Information). As discussed below, we were able to further divide the NR-PKSs into seven groups based on phylogeny, domain structure, and known products (Figure 2 and Supplementary Table S1 and Figure S3, Supporting Information).

A great deal of work by several laboratories over many years[19] has revealed the chemical products of six *A. nidulans* NR-PKSs (StcA,[20,21] WA,[22] MdpG,[6,10,14] AptA,[23,24] OrsA,[6,8,25] and AfoE[11]) (Figure 2). To determine the products of the remaining eight NR-PKSs, we replaced their native promoters with the *alcA* promoter[18] (resulting strains are listed in Supplementary Table S3, Supporting Information). We carried out the replacements in strains in which the production of the major natural product, sterigmatocystin, was eliminated.[6] This simplifies analysis, and we anticipated that in some cases, elimination of sterigmatocystin production would free up subunits for incorporation into other NPs. Strains were grown in inducing conditions, and natural products were extracted from the medium, or mycelia, and were subjected to HPLC-DAD-MS metabolite profiling (Figure 3). Gratifyingly, replacement of the promoters of five NR-PKS genes (ANID_06448.1, ANID_08383.1, ANID_00523.1, ANID_07903.1, and ANID_03230.1 using the Broad Institute gene designations (http://www.broadinstitute.org/annotation/genome/aspergillus_group/MultiHome.html) resulted in the production of at least one major compound upon induction of the *alcA* promoter. We designate these genes *pkbA* (ANID_06448.1), *pkdA* (ANID_00523.1), *pkeA* (ANID_07903.1), and *pkfA* (ANID_03230.1). The gene ANID_08383.1 has been designated *ausA* previously.[26,27] All the major polyketides produced were purified from large-scale cultures, and the structures were elucidated by spectroscopic methods (Supplementary Figure S6, Supporting Information; for details on structural elucidation, see Supporting Information). These compounds are listed in Figure 2 (compounds **1**–**7**).

Induction of ANID_07071.1 resulted in the production of several NPs but all in small amounts (Figure 3). ANID_07071.1 does not contain a releasing domain, and on the basis of previous data,[28] we hypothesized that another gene in the cluster is required for product release. ANID_07070.1, which is adjacent to ANID_07071.1, encodes a β-lactamase-type thioesterase that potentially might be involved in product release. We created a strain in which both the promoters of ANID_07071.1 and ANID_07070.1 were replaced with the *alcA* promoter. As hypothesized, induction of this strain resulted in a dramatically increased yield (Figure 3). We designate ANID_07071.1 *pkgA* and ANID_07070.1 *pkgB*. A larger scale induction of an ANID_07071.1 and ANID_07070.1 promoter replacement strain followed by extensive chromatography led to the isolation of two major compounds, alternariol (**8**),[15] an important toxin with antifungal, phytotoxic, and anticholinesterase activity previously reported from *Alternaria sp.*, and dehydrocitreoisocoumarin (**9**), as well as four minor compounds including two heptaketide isocoumarins (**10** and **11**) and two hexaketide isocoumarins (**12** and **13**). It is interesting to note that ANID_07071.1 and WA both are heptaketide synthases. Instead of the intramolecular Claisen

dx.doi.org/10.1021/ja3016395 | *J. Am. Chem. Soc.* 2012, 134, 8212–8221

cyclization that occurs in WA, which has a functional TE/CLC domain,[29] with ANID_07071.1, which lacks a TE/CLC domain, the released heptaketide undergoes lactonization to produce **9**, which converts to **8** *in vivo* (Figure 2).

Group and domain architecture	NR-PKS gene and mode of cyclization	Starter unit	Substrate incorporated	Release mechanism	Product released from NR-PKS	Downstream metabolites	References
Group I	ANID_07909.1 (OrsA) C2-C7		3 × Mal-CoA	TE	orsellinic acid	F9775A and B, sanghaspirodins A and B.	6, 8, 25, 43
Group III	ANID_08209.1 (WA) C2-C7		6 × Mal-CoA	TE/CLC	2-acetoacetyl T4HN	YWA1 and melanin.	22
Group IV	ANID_07825.1 (StcA) C4-C9, C2-C11	StcJ StcK	7 × Mal-CoA	TE/CLC	norsolorinic acid anthrone	sterigmatocystin.	20, 21
Group V	ANID_06000.1 (AptA) C6-C11, C4-C13		8 × Mal-CoA	AptC + β-TE (AptB)	2,3,6,8,9-pentahydroxy-1-oxo-3-(2-oxopropyl)-1,2,3,4-tetrahydroanthracene-2-carboxylic acid	asperthecin.	23, 24
	ANID_00150.1 (MdpG) C6-C11, C4-C13		7 × Mal-CoA	β-TE (MdpF)	atrochrysone carboxylic acid	emodin and monodictylphenone, phenyl xanthones, sanghaspirodins A and B.	6, 10, 14, 43
	ANID_07071.1 (PkgA) C2-C7		6 × Mal-CoA	β-TE (PkgB)	dehydrocitreoisocoumarin (**9**); host reductase; citreoisocoumarin (**10**); 6,8-dihydroxy-3-(4-hydroxy-2-oxopentyl)-isocoumarin (**11**); $-H_2O$; alternariol (**8**)		This study
			5 × Mal-CoA	β-TE (PkgB)	6,8-dihydroxy-3-(2-oxopropyl)-isocoumarin (**12**); host reductase; 6,8-dihydroxy-3-(2-hydroxy-propyl)-isocoumarin (**13**)		This study
Group VI	ANID_06448.1 (PkbA) C2-C7		3 × Mal-CoA + 1 × SAM	TE	3-methylorsellinic acid (**1**); $-CO_2$; 2,5-dimethylresorcinol (**2**); cichorine (**19**)		This study
	ANID_08383.1 (AusA) C2-C7		3 × Mal-CoA + 2 × SAM	TE	3,5-dimethylorsellinic acid (**3**)	austinol and dehydroaustinol.	26, 27, and this study

Figure 2. continued

dx.doi.org/10.1021/ja3016395 | J. Am. Chem. Soc. 2012, 134, 8212−8221

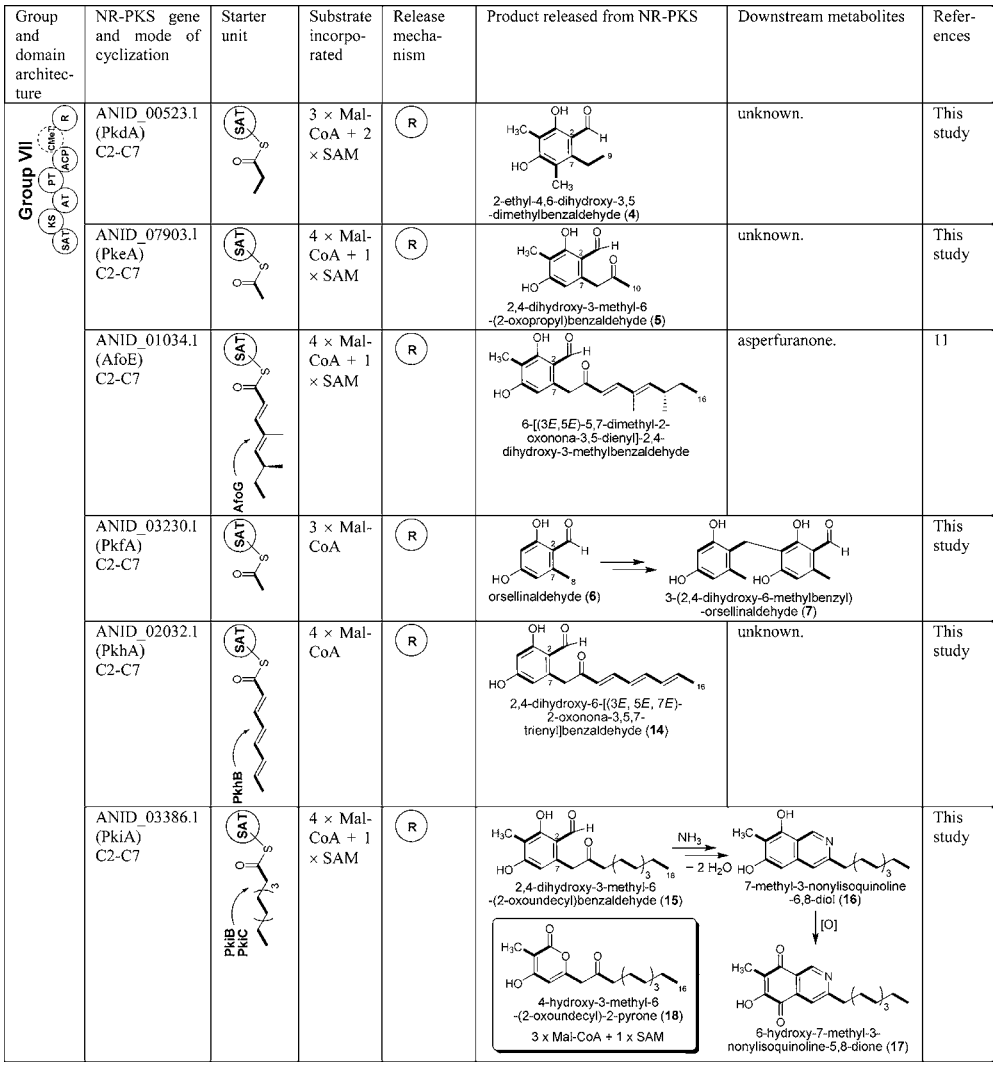


Figure 2. The NR-PKSs of *A. nidulans*, their biosynthetic characteristics, and their products. Domain abbreviations: SAT = starter unit-ACP transacylase, KS = ketosynthase, AT = acyl transferase, PT = product template, ACP = acyl carrier protein, TE = thioesterase, TE/CLC = thioesterase/Claisen cyclase, CMeT = C-methyltransferase, R = reductase, β-TE = β-lactamase-type thioesterase. With respect to domain structure, domains encircled with a dotted line are present in some cases and not others. With respect to starter unit, it has been proposed that the cysteine or serine in the GXCXG or GXSXG motif in the SAT domain is responsible for transferring the starter unit by using thioester or oxyester chemistry, respectively.[44] The GXGXG motif in the SAT domain might be inactive or have an unknown function. Most NR-PKSs identified from the *A. nidulans* genome have the GXCXG motif except for AptA and AN07071.1, which have a GXSXG motif, and MdpG and OrsA, which have a GXGXG motif in their SAT domains. Release mechanisms: TE cleaves the thioester bond and releases the acid product; TE/CLC releases the product without the acid functional group via intramolecular Claisen cyclization; β-TE cleaves the thioester bond and releases the acid product but cleavage is due to a protein acting in trans, not to a domain of the NR-PKS (the protein acting in trans is listed in each case); R cleaves the product via reductive release such that the product has an aldehyde group. With respect to downstream products, for previously published studies only the final product of the pathway is listed. StcJ (ANID_07815.1) and StcK (ANID_07814.1) and PkiB (ANID_03380.1) and PkiC (ANID_03381.1) are FASs that synthesize starter units for StcA and PkiA, respectively. AfoG (ANID_01036.1) and PkhB (ANID_02035.1) are HR-PKSs that synthesize starter units for AfoE and PkhA, respectively. AptB (ANID_06001.1), MdpF (ANID_00149.1), and PkgB (ANID_07070.1) are β-TEs near the NR-PKSs AptA, MdpG, and PkgA, respectively. AptC (ANID_06002.1) is a C2-hydroxylase in the Apt cluster.[24] Note: Induction of ANID_07071.1 resulted in production of compound **9** - **13**. However, compounds **12** and **13** were found in small amounts. These compounds result from incorporation of five Mal-CoAs, and they indicate that although ANID_07071.1 is a heptaketide synthase, it is not strict with respect to the number of Mal-CoA molecules incorporated. Compound **18** is a self-released shunt product after incorporating three Mal-CoAs and one SAM.

The fact that induction of the two remaining NR-PKS genes, ANID_02032.1 and ANID_03386.1, did not produce compounds suggested that coexpression of nearby genes might be necessary to provide specialized starter units for these NR-PKSs.[11,30] ANID_02032.1 is located next to an HR-PKS, ANID_02035.1, and ANID_03386.1 is located near ANID_03380.1 and ANID_03381.1, putative α and β subunits of a fatty acid synthase (FAS). We created a strain where both the promoters of the NR-PKS ANID_02032.1 and the HR-PKS ANID_02035.1 were replaced with the *alcA* promoter. The resulting strain produced at least two UV−vis detectable compounds upon induction, although in very small amounts (Figure 3). We designate ANID_02032.1 *pkhA* and ANID_02035.1 *pkhB*. We were able to isolate one major compound, 2,4-dihydroxy-6-[(3*E*,5*E*,7*E*)-2-oxonona-3,5,7-trienyl]benzaldehyde (**14**) (Figure 2), from the scale-up culture

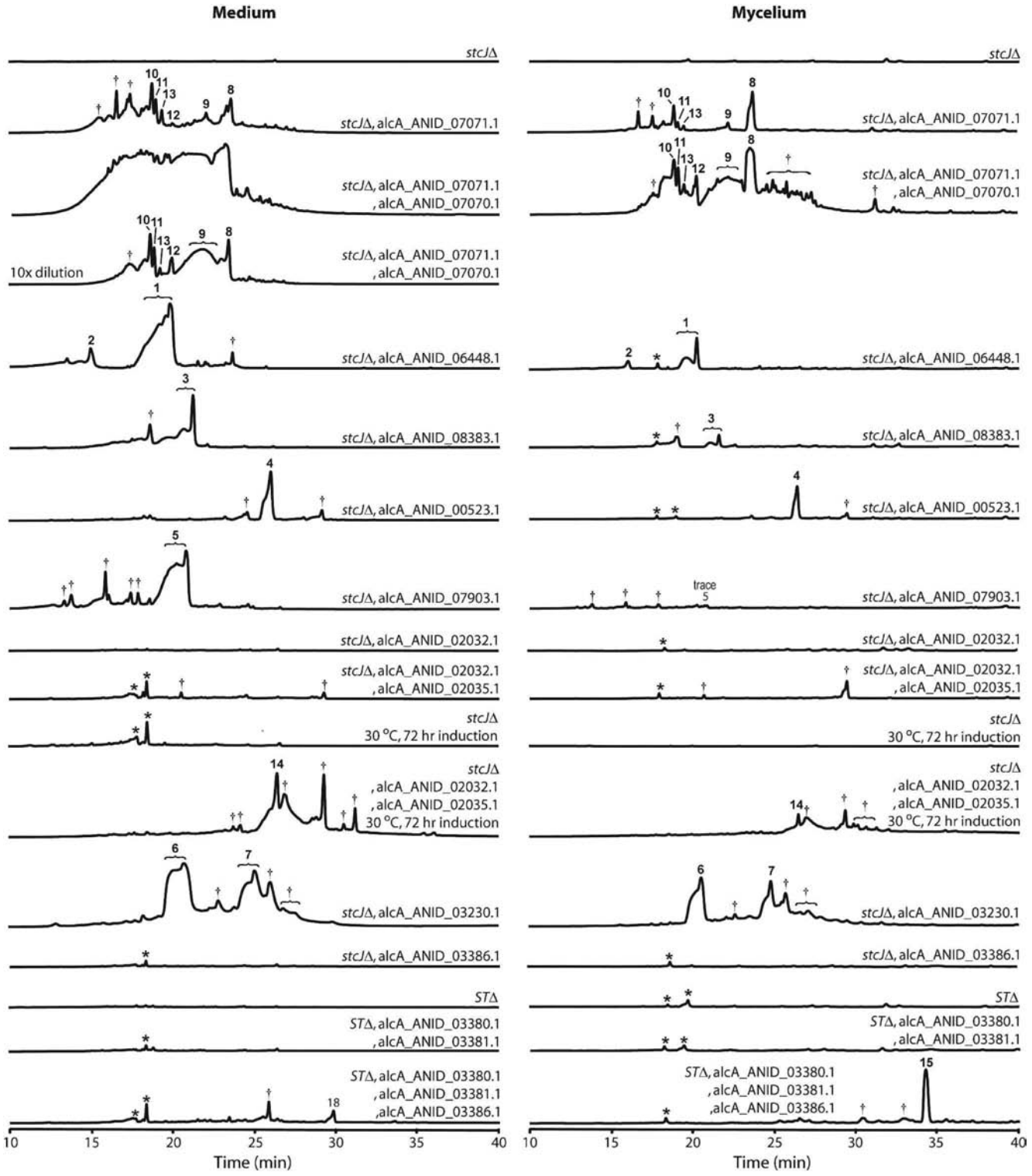


Figure 3. Total scan HPLC profiles of natural products extracted from the culture medium and mycelium of *alcA* promoter-replaced strains. Peaks that are also present in the *stcJ*Δ control strain are indicated with an asterisk (∗) and are considered not to be specific to the NR-PKS activation. Peaks that are specifically present in induced NR-PKS promoter-replacement strains but are not structurally elucidated due to low yield or poor stability are marked with a dagger (†).

when we lowered the induction temperature from 37 to 30 °C and extended the incubation time from 48 to 72 h (Figure 3). Our data thus indicate that the HR-PKS, ANID_02035.1, produces the octatrienoyl starter that is loaded onto the SAT domain of the NR-PKS, ANID_02032.1 (Figure 2).

We created a strain where the promoters of ANID_03386.1 and the two FAS subunit genes ANID_03380.1 and ANID_03381.1 were all replaced with the *alcA* promoter. The two FAS subunits are divergently transcribed, and we developed a procedure to replace the promoters of both genes

dx.doi.org/10.1021/ja3016395 | J. Am. Chem. Soc. 2012, 134, 8212−8221

using a single selectable marker (Figure 1B). Induced expression of all three genes led to the isolation of a major product, 2,4-dihydroxy-3-methyl-6-(2-oxoundecyl)benzaldehyde (**15**), from the mycelium (Figure 3). We also isolated compounds **16** and **17** from the mycelium and **18** from the medium (Figure 2). We designate ANID_03386.1 *pkiA*, ANID_03380.1 *pkiB*, and ANID_03381.1 *pkiC*. These data indicate that a decanoyl starter unit is first synthesized by the FAS (ANID_03380.1 and ANID_03381.1) and then loaded onto ANID_03386.1. Since we did not detect any decanoic acid, the decanoyl starter is likely to be transferred directly from the FAS to the SAT domain of ANID_03386.1, similar to the biosynthesis of norsolorinic acid anthrone.[31] Compound **15** is then transaminated and aromatized to become **16**, which is then oxidized to generate **17**. The minor compound **18**, thus, is a shunt product (Figure 2).

Based on these data, we were able to determine the PKSs required for the production of three *A. nidulans* natural products, a process that is often difficult and time-consuming. A labeled precursor approach[32] has revealed that 3,5-dimethylorsellinic acid (**3**) (Figure 2) is the precursor of austinol and dehydroaustinol, two meroterpenoids isolated from *A. nidulans*.[23] The fact that 3,5-dimethylorsellinic acid (**3**) is the product of ANID_08383.1 suggested that this PKS is required for the synthesis of these meroterpenoids. This has recently been confirmed by another group and by us.[26,27] We have recently isolated the phytotoxin cichorine (**19**) (Figure 2) from *A. nidulans*.[14] The chemical structure of cichorine suggested it might be synthesized by modification of the product of ANID_06448.1, that is, compound **1**. We deleted ANID_06448.1 and found that cichorine production was, indeed, eliminated (Supplementary Figure S7, Supporting Information). ANID_06448.1 is, thus, the PKS responsible for cichorine biosynthesis.

Phylogenetic Relationships of NR-PKSs and Carbon Skeletons of Their Products. The genus *Aspergillus* contains more than 150 species, and many of them are potent producers of NP.[33] The sequencing of the genomes of members of this genus presents an opportunity to tap this potentially rich source of important compounds, but a key missing element for the many species without developed molecular genetics systems is the ability to correlate NP gene clusters with their products easily. For example, if a species with a sequenced genome but no molecular genetic system was found to produce a potentially useful compound but only at low levels, it would be difficult to exploit the potential utility of the compound. The ability to predict which PKS is essential for the production of the product might allow one to identify the gene cluster responsible for the production of the product, and this cluster could be cloned and expressed heterologously to produce the desired compound in larger amounts.

Our data greatly expand the knowledge of polyketides produced by NR-PKSs in *A. nidulans*, and we wondered whether our data in combination with sequence information might allow us to make useful predictions about the products of NR-PKSs in other *Aspergillus* species. We reasoned that a phylogenetic analysis of NR-PKSs might be useful because such an analysis includes key information on the presence, absence, and order of functional NR-PKS domains. We used a previous phylogenetic analysis of KS domains by Kroken et al. as a starting point for our analyses.[34] The analysis of Kroken et al. classified NR-PKSs into three subclades. Subclade I and II NR-PKSs do not contain a methyltransferase (CMeT) domain and are associated with aflatoxin and melanin biosynthesis, respectively. It should be noted that when this initial analysis was made, no product was associated with any of the NR-PKS in subclade III, which are NR-PKSs that contain CMeT domains. Since the product template (PT) domain embedded in the middle of the NR-PKSs has been shown to control regioselective cyclization,[35−37] Li et al. further categorized subclade I and II NR-PKSs into five major groups with each group corresponding to a unique product size or cyclization regioselectivity using PT phylogeny.[38]

In this study, we have identified the products of eight NR-PKSs in *A. nidulans* and consequently have greatly increased the number of fungal NR-PKSs for which the products are known. In particular, we have increased the number of NR-PKSs in subclade III with associated products, and this has now allowed us to further classify subclade III into groups VI and VII (Supplementary Figure S3, Supporting Information). All group VI and VII NR-PKSs analyzed so far produce monocyclic aromatics with C2−C7 cyclization regioselectivity (Figure 2). Noticeably, some CMeT domain-less NR-PKSs like ANID_03230.1 and ANID_02032.1 fell into group VII. These NR-PKSs produce monocyclic aromatics without methyl branching on their benzene rings (Figure 2). We further performed phylogenetic analysis of PT domains extracted from the known NR-PKSs and found, interestingly, that the phylogeny obtained using the PT domains is similar to the phylogeny obtained using simply the full length NR-PKS (Supplementary Table S1 and Figures S3, and S4, Supporting Information). Our analysis showed that NR-PKSs within the same group produce similar polyketides except NR-PKSs belonging to group V, which lacks the product-releasing domain. NR-PKSs in group V seem to produce the most diverse aromatic polyketides. They require separate genes to release their products, and they generate multicyclic aromatics with various chain lengths and cyclization regioselectivities. The known NR-PKSs in group V include hepta- (ANID_07071.1 and GsfA[39]), octa- (ACAS[28] and MdpG[10]), nona- (AptA[24] and VrtA[39]), and decaketide synthases (AdaA[24]) with C2−C7 (ANID_07071.1), C6−C1 (GsfA), and C6−C11 (ACAS, MdpG, AptA, VrtA, and AdaA) cyclization modes.

To determine whether the phylogenetic analysis that we have done in *A. nidulans* could be applied more broadly to *Aspergillus* NR-PKSs sequenced so far, we analyzed 71 fungal NR-PKSs available at the Broad Institute *Aspergillus* Comparative Database (Supplementary Table S2, Supporting Information, and Figure 4). Gratifyingly, 69 of 71 NR-PKSs from the *Aspergillus* Comparative Database fell into these seven groups. The only two outliers are AFL2G_04689 and AO090023000877, which are more similar to MSAS, a PR-PKS from *A. terreus*.[40] Among 51 unknown NR-PKSs analyzed, 48 fell into groups III−VII, and the previously underrepresented groups VI and VII contained the greatest number of NR-PKSs. Although currently it is still not possible to predict the exact product of the NR-PKS using bioinformatics analysis alone, our data shows that at least for *Aspergillus* species genomes sequenced so far, 69 of 71 NR-PKSs with canonical domain architectures can be grouped into these seven groups and, with the exception of group V, the polyketide products of the PKSs within each group are structurally similar. This is an important step for genome mining of aromatic polyketide biosynthesis genes in *Aspergillus* species with sequenced genomes because if a valuable compound is identified, the structure of the compound will, in many cases, narrow the

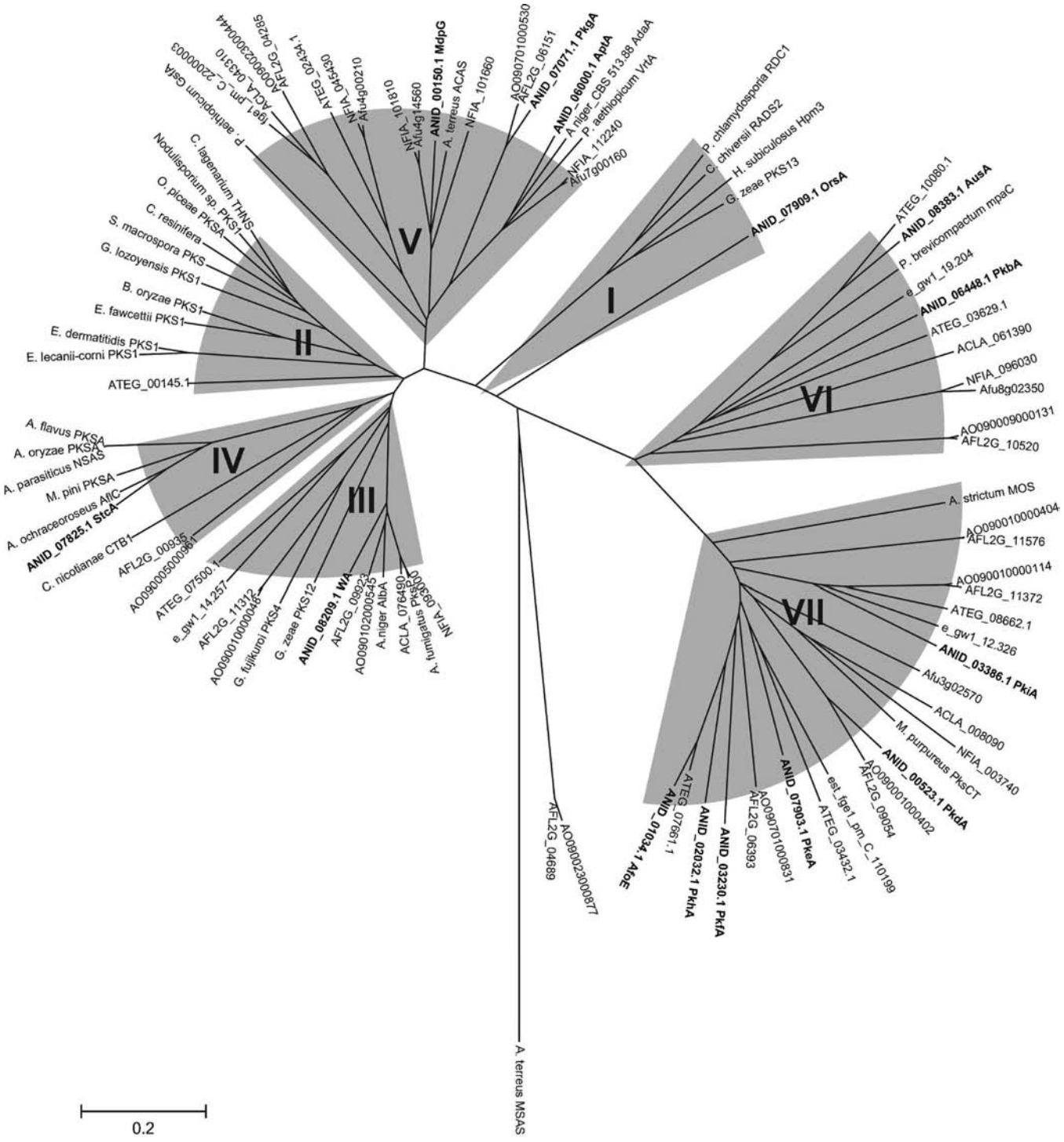


Figure 4. Phylogenetic analysis of 71 NR-PKSs obtained from the Broad Institute *Aspergillus* Comparative Database (http://www.broadinstitute.org/annotation/genome/aspergillus_group/MultiHome.html). NR-PKSs from *A. nidulans* are labeled in boldface. Domain architectures of protein sequences are listed in Table S3, Supporting Information. *A. terreus* MSAS, a PR-PKS, was used as the outgroup. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree.

search for NR-PKSs responsible for its production into a particular phylogenetic group containing only a few members.

■ DISCUSSION

The approach that we have developed has allowed us to complete the determination of the products of aromatic polyketide synthases in *A. nidulans*, including seven novel compounds (**4**, **7**, **14**, **15**, **16**, **17**, and **18**). This represents a major step in exploiting the diversity of fungal natural products coded in genomes. Several important conclusions can be drawn from our results. First, it is feasible and practical to investigate *A. nidulans* natural products by bypassing normal regulation. Indeed, much more time and effort was expended in analyzing products of the expression strains than in creating them. Our efforts have resulted in the discovery that *A. nidulans* can produce the important toxin alternariol (**8**), the identification

dx.doi.org/10.1021/ja3016395 | J. Am. Chem. Soc. 2012, 134, 8212−8221

of the PKS responsible for its synthesis, and the identification of the PKS responsible for cichorine (**19**) biosynthesis. Second, we have shown that serial promoter replacements are feasible and not prohibitively time-consuming and, thus, that there are no conceptual barriers to up-regulating all the genes of target clusters. Serial promoter replacement should allow the overproduction of the final products of target natural product gene clusters (when the promoters of all genes in the cluster are replaced) as well as intermediates in the pathways (when some of the genes in the cluster are replaced). Replacing the promoters of all genes in large clusters would require the development of additional selectable markers or techniques to recycle markers, but this is certainly feasible. This promises to be valuable in translating NPs into products. Although we have focused initially on NR-PKSs, the approach should work for other classes of NP biosynthetic pathways and should be applicable to other fungal species with developed molecular genetic systems. Third, our data validate our previous suggestion that if an NR-PKS does not contain a releasing domain, a separate gene in the same cluster encoding a β-lactamase thioesterase is required for releasing the compound.[23,28] This is true for three NR-PKSs in *A. nidulans*, and 23 of the 71 NR-PKSs in the *Aspergillus* species genomes sequenced so far lack product releasing domains. Seventeen of them have a β-lactamase thioesterase or an esterase gene nearby (Supplementary Table S2, Supporting Information). Release of NR-PKS products by β-lactamase thioesterase genes separate from the NR-PKS thus appears to be common in *Aspergillus* species. Fourth, our data indicate that if an HR-PKS or FAS is present in a cluster with an NR-PKS, it is very likely to provide the starter unit loaded onto the SAT domain of the NR-PKS. This is true for *A. nidulans* NR-PKSs, and 17 of the 71 NR-PKSs in the *Aspergillus* species genomes sequenced so far are in clusters with an HR-PKS or FAS (Supplementary Table S2, Supporting Information). Production of starter units for NR-PKSs by HR-PKSs or FASs thus appears to be a frequent feature of fungal secondary metabolism. Fifth, our data in combination with previous data[19,24] indicate that each of the 14 NR-PKSs in *A. nidulans* produces a unique product (Figure 2). This reveals that the PKSs themselves generate a great deal of the diversity of fungal natural products. This diversity is then multiplied by the enzymes that modify the PKS products resulting, in principle, in huge numbers of different NPs. Finally, our results have greatly expanded the data on the products generated by fungal NR-PKSs and will facilitate genome mining efforts by narrowing down the number of target genes that need to be experimentally verified.

In view of previous data,[6,9−11] we were surprised that up-regulation of transcription factors generally yielded no or inadequate increases in compound production. We do not know why this was the case. It could be due to post-translational down-regulation of the activity of the transcription factor as has been demonstrated for the AflR transcription factor,[41] to other mechanisms of transcriptional inhibition as has been demonstrated in *A. terreus*,[42] or to other, as yet unknown, mechanisms. In any case, simple up-regulation of transcription factors associated with SM clusters does not appear to be a generally successful strategy for obtaining expression of SM gene clusters.

CONCLUSION

We have developed a strategy that allows us to obtain expression of cryptic NP genes by bypassing normal regulatory mechanisms. Using this strategy, we have successfully deciphered the products of eight NR-PKSs in *A. nidulans*. Furthermore, structural information derived from this work has allowed improved prediction of the carbon skeleton of aromatic polyketides from NR-PKS sequences through structure−phylogenetic analysis. This will greatly facilitate the elucidation of NR-PKS biosynthetic pathways in other fungal species.

EXPERIMENTAL SECTION

Molecular Genetic Manipulations. Replacement of endogenous promoters with the *alcA* promoter was carried out as shown in Figure 1. Primers used in this study are listed in Supplemental Table S4, Supporting Information. Transforming fragments were generated by fusion PCR as described[16,17] except that in most cases KOD DNA polymerase (EMD Biosciences) was used instead of Accuprime Taq HiFi because of the stronger proofreading activity of KOD. With KOD enzyme, fusions were carried out at T_m (for the lowest melting temperature primer) + 2 °C and the extension time was 30 s per kilobase of expected fusion product. LO2026 (Supplemental Table S3, Supporting Information) was used as a recipient strain for most transformations, and the *Aspergillus fumigatus pyrG* (*AfpyrG*) and *pyroA* (*AfpyroA*) genes were used as selectable markers. For the ANID_03380.1, ANID_03381.1, ANID_03386.1 gene cluster, LO4389 was used as the recipient strain (Supplemental Table S3, Supporting Information). Replacing the promoters of ANID_03380.1 and ANID_03381.1 presented a problem because the two genes (which encode fatty acid synthase subunits) are divergently transcribed and the distance between the two coding sequences is only about 600 bp. With our normal promoter replacement procedure, replacement of the promoter of one of the two genes with the *alcA* promoter would have been straightforward but subsequent replacement of the promoter of the second gene would have resulted in the deletion of the *alcA* promoter of the first gene. To circumvent this problem, we developed the procedure shown in Figure 1B. Fusion PCR was used to create two transforming fragments. One contained a portion of ANID_03381.1 fused to the *alcA* promoter and a fragment carrying a portion of the *AfpyrG* gene. The second contained a portion of the *AfpyrG* gene and the *alcA* promoter fused to ANID_03380.1. The two fragments of *AfpyrG* overlapped by 540 bp. Transformation resulted in ANID_03381.1 and ANID_03380.1 being driven by separate copies of the *alcA* promoter and fusion of the two portions of *AfpyrG* creating a full-length, functional copy. The ANID_06448.1 deletion strain was generated by replacing each gene with the *Aspergillus fumigatus pyrG* gene in the *A. nidulans* strain LO2026 (*nkuA*Δ, *stcJ*Δ).[6] The construction of fusion PCR products, protoplast production, and transformation were carried out as described previously,[16,17] For the construction of the fusion PCR fragments, two 1000-bp fragments of genomic *A. nidulans* DNA, upstream and downstream of the targeted gene, were amplified by PCR. All transformants were verified by diagnostic PCR (Supplemental Figure S5, Supporting Information). Genotypes of all strains are given in Table S3 in the Supporting Information.

Fermentation and LC/MS Analysis. For *alcA*(p) induction, 30×10^6 spores were grown in 30 mL of liquid LMM medium (15 g/L lactose, 6 g/L $NaNO_3$, 0.52 g/L KCl, 0.52 g/L $MgSO_4{\cdot}7H_2O$, 1.52 g/L KH_2PO_4, 1 mL/L trace element) in 125 mL flasks at 37 °C with shaking at 180 rpm and supplemented with uracil (1 g/L), uridine (10 mM), riboflavin (2.5 mg/L), or pyridoxine (0.5 mg/L) when necessary. Cyclopentanone at a final concentration of 10 mM was added to the medium 18 h after inoculation. Culture medium was collected 48 h after cyclopentanone induction by filtration and extracted with the same volume of EtOAc twice. The mycelium collected was soaked in 50 mL of MeOH for one day. After removal of the cell debris by filtration, MeOH was collected, concentrated, resuspended in 25 mL of ddH_2O, and extracted with the same volume of EtOAc twice. EtOAc from the combined EtOAc layers was evaporated by TurboVap LV (Caliper LifeSciences). The crude extracts were then redissolved in 0.5 mL of DMSO/MeOH (1:4), and 10 μL was injected for LC-DAD-MS analysis as described previously.[6]

dx.doi.org/10.1021/ja3016395 | J. Am. Chem. Soc. 2012, 134, 8212−8221

■ ASSOCIATED CONTENT

Supporting Information

Details for large scale purifications of compounds from each strain, full characterization of new compounds, primer sequences, diagnostic PCR, and phylogenetic analyses. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

ymchiang@mail.chna.edu.tw; boakley@ku.edu; clayw@usc.edu

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

Research in the Wang and Oakley Groups is supported in part by PO1-GM084077 from the National Institute of General Medical Sciences. The Oakley group is additionally supported by the University of Kansas Endowment Fund.

■ REFERENCES

(1) Galagan, J. E.; Calvo, S. E.; Cuomo, C.; Ma, L. J.; Wortman, J. R.; Batzoglou, S.; Lee, S. I.; Basturkmen, M.; Spevak, C. C.; Clutterbuck, J.; Kapitonov, V.; Jurka, J.; Scazzocchio, C.; Farman, M.; Butler, J.; Purcell, S.; Harris, S.; Braus, G. H.; Draht, O.; Busch, S.; D'Enfert, C.; Bouchier, C.; Goldman, G. H.; Bell-Pedersen, D.; Griffiths-Jones, S.; Doonan, J. H.; Yu, J.; Vienken, K.; Pain, A.; Freitag, M.; Selker, E. U.; Archer, D. B.; Penalva, M. A.; Oakley, B. R.; Momany, M.; Tanaka, T.; Kumagai, T.; Asai, K.; Machida, M.; Nierman, W. C.; Denning, D. W.; Caddick, M.; Hynes, M.; Paoletti, M.; Fischer, R.; Miller, B.; Dyer, P.; Sachs, M. S.; Osmani, S. A.; Birren, B. W. *Nature* **2005**, *438*, 1105−1115.

(2) Keller, N. P.; Turner, G.; Bennett, J. W. *Nat. Rev. Microbiol.* **2005**, *3*, 937−947.

(3) Chiang, Y. M.; Lee, K. H.; Sanchez, J. F.; Keller, N. P.; Wang, C. C. *Nat. Prod. Commun.* **2009**, *4*, 1505−1510.

(4) Scherlach, K.; Hertweck, C. *Org. Biomol. Chem.* **2009**, *7*, 1753−1760.

(5) Chiang, Y. M.; Chang, S. L.; Oakley, B. R.; Wang, C. C. *Curr. Opin. Chem. Biol.* **2011**, *15*, 137−143.

(6) Bok, J. W.; Chiang, Y. M.; Szewczyk, E.; Reyes-Domingez, Y.; Davidson, A. D.; Sanchez, J. F.; Lo, H. C.; Watanabe, K.; Strauss, J.; Oakley, B. R.; Wang, C. C.; Keller, N. P. *Nat. Chem. Biol.* **2009**, *5*, 462−464.

(7) Bode, H. B.; Bethe, B.; Hofs, R.; Zeeck, A. *ChemBioChem* **2002**, *3*, 619−627.

(8) Schroeckh, V.; Scherlach, K.; Nutzmann, H. W.; Shelest, E.; Schmidt-Heck, W.; Schuemann, J.; Martin, K.; Hertweck, C.; Brakhage, A. A. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 14558−14563.

(9) Bergmann, S.; Schumann, J.; Scherlach, K.; Lange, C.; Brakhage, A. A.; Hertweck, C. *Nat. Chem. Biol.* **2007**, *3*, 213−217.

(10) Chiang, Y. M.; Szewczyk, E.; Davidson, A. D.; Entwistle, R.; Keller, N. P.; Wang, C. C.; Oakley, B. R. *Appl. Environ. Microbiol.* **2010**, *76*, 2067−2074.

(11) Chiang, Y. M.; Szewczyk, E.; Davidson, A. D.; Keller, N.; Oakley, B. R.; Wang, C. C. *J. Am. Chem. Soc.* **2009**, *131*, 2965−2970.

(12) Cox, R. J. *Org. Biomol. Chem.* **2007**, *5*, 2010−2026.

(13) Crawford, J. M.; Townsend, C. A. *Nat. Rev. Microbiol.* **2010**, *8*, 879−889.

(14) Sanchez, J. F.; Entwistle, R.; Hung, J. H.; Yaegashi, J.; Jain, S.; Chiang, Y. M.; Wang, C. C.; Oakley, B. R. *J. Am. Chem. Soc.* **2011**, *133*, 4010−4017.

(15) Cole, R. J., Cox, R. H., Eds. *Handbook of Toxic Fungal Metabolites*; Academic Press: New York, 1981; p 614.

(16) Nayak, T.; Szewczyk, E.; Oakley, C. E.; Osmani, A.; Ukil, L.; Murray, S. L.; Hynes, M. J.; Osmani, S. A.; Oakley, B. R. *Genetics* **2006**, *172*, 1557−1566.

(17) Szewczyk, E.; Nayak, T.; Oakley, C. E.; Edgerton, H.; Xiong, Y.; Taheri-Talesh, N.; Osmani, S. A.; Oakley, B. R. *Nat. Protoc.* **2006**, *1*, 3111−3120.

(18) Felenbok, B. *J. Biotechnol.* **1991**, *17*, 11−17.

(19) Chiang, Y. M.; Oakley, B. R.; Keller, N. P.; Wang, C. C. *Appl. Microbiol. Biotechnol.* **2010**, *86*, 1719−1736.

(20) Brown, D. W.; Yu, J. H.; Kelkar, H. S.; Fernandes, M.; Nesbitt, T. C.; Keller, N. P.; Adams, T. H.; Leonard, T. J. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 1418−1422.

(21) Yu, J. H.; Leonard, T. J. *J. Bacteriol.* **1995**, *177*, 4792−4800.

(22) Watanabe, A.; Fujii, I.; Sankawa, U.; Mayorga, M. E.; Timberlake, W. E.; Ebizuka, Y. *Tetrahedron Lett.* **1999**, *40*, 91−94.

(23) Szewczyk, E.; Chiang, Y. M.; Oakley, C. E.; Davidson, A. D.; Wang, C. C.; Oakley, B. R. *Appl. Environ. Microbiol.* **2008**, *74*, 7607−7612.

(24) Li, Y.; Chooi, Y. H.; Sheng, Y.; Valentine, J. S.; Tang, Y. *J. Am. Chem. Soc.* **2011**, *133*, 15773−15785.

(25) Sanchez, J. F.; Chiang, Y. M.; Szewczyk, E.; Davidson, A. D.; Ahuja, M.; Elizabeth Oakley, C.; Woo Bok, J.; Keller, N.; Oakley, B. R.; Wang, C. C. *Mol. Biosyst.* **2010**, *6*, 587−593.

(26) Nielsen, M. L.; Nielsen, J. B.; Rank, C.; Klejnstrup, M. L.; Holm, D. M.; Brogaard, K. H.; Hansen, B. G.; Frisvad, J. C.; Larsen, T. O.; Mortensen, U. H. *FEMS Microbiol. Lett.* **2011**, *321*, 157−166.

(27) Lo, H. C.; Entwistle, R.; Guo, C. J.; Ahuja, M.; Szewczyk, E.; Hung, J. H.; Chiang, Y. M.; Oakley, B. R.; Wang, C. C. *J. Am. Chem. Soc.* **2012**, *134*, 4709−4720.

(28) Awakawa, T.; Yokota, K.; Funa, N.; Doi, F.; Mori, N.; Watanabe, H.; Horinouchi, S. *Chem. Biol.* **2009**, *16*, 613−623.

(29) Fujii, I.; Watanabe, A.; Sankawa, U.; Ebizuka, Y. *Chem. Biol.* **2001**, *8*, 189−197.

(30) Crawford, J. M.; Dancy, B. C.; Hill, E. A.; Udwary, D. W.; Townsend, C. A. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 16728−16733.

(31) Watanabe, C. M.; Townsend, C. A. *Chem. Biol.* **2002**, *9*, 981−988.

(32) Simpson, T. J.; Stenzel, D. J.; Bartleet, A. J.; O'Brien, E.; Holker, J. S. E. *J. Chem. Soc., Perkin Trans. 1* **1982**, 2687−2692.

(33) Geiser, D. M.; Klich, M. A.; Frisvad, J. C.; Peterson, S. W.; Varga, J.; Samson, R. A. *Stud. Mycol.* **2007**, *59*, 1−10.

(34) Kroken, S.; Glass, N. L.; Taylor, J. W.; Yoder, O. C.; Turgeon, B. G. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 15670−15675.

(35) Crawford, J. M.; Thomas, P. M.; Scheerer, J. R.; Vagstad, A. L.; Kelleher, N. L.; Townsend, C. A. *Science* **2008**, *320*, 243−246.

(36) Crawford, J. M.; Korman, T. P.; Labonte, J. W.; Vagstad, A. L.; Hill, E. A.; Kamari-Bidkorpeh, O.; Tsai, S. C.; Townsend, C. A. *Nature* **2009**, *461*, 1139−1143.

(37) Liu, T.; Chiang, Y. M.; Somoza, A. D.; Oakley, B. R.; Wang, C. C. *J. Am. Chem. Soc.* **2011**, *133*, 13314−13316.

(38) Li, Y.; Image, I. I.; Xu, W.; Image, I.; Tang, Y. *J. Biol. Chem.* **2010**, *285*, 22764−22773.

(39) Chooi, Y. H.; Cacho, R.; Tang, Y. *Chem. Biol.* **2010**, *17*, 483−494.

(40) Fujii, I.; Ono, Y.; Tada, H.; Gomi, K.; Ebizuka, Y.; Sankawa, U. *Mol. Gen. Genet.* **1996**, *253*, 1−10.

(41) Shimizu, K.; Hicks, J. K.; Huang, T. P.; Keller, N. P. *Genetics* **2003**, *165*, 1095−1104.

(42) Gressler, M.; Zaehle, C.; Scherlach, K.; Hertweck, C.; Brock, M. *Chem. Biol.* **2011**, *18*, 198−209.

(43) Scherlach, K.; Sarkar, A.; Schroeckh, V.; Dahse, H. M.; Roth, M.; Brakhage, A. A.; Horn, U.; Hertweck, C. *ChemBioChem* **2011**, *12*, 1836−1839.

(44) Crawford, J. M.; Vagstad, A. L.; Ehrlich, K. C.; Townsend, C. A. *Bioorg. Chem.* **2008**, *36*, 16−22.