ACS SyntheticBiology

Research Article

pubs.acs.org/synthbio

Native Promoter Strategy for High-Yielding Synthesis and Engineering of Fungal Secondary Metabolites

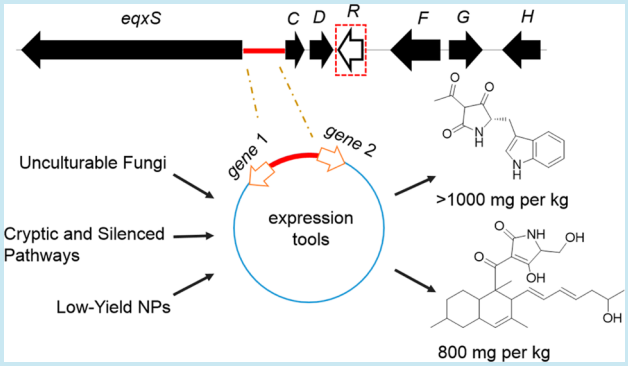
Thomas B. Kakule,[†] Raquel C. Jadulco,[‡] Michael Koch,[‡] Jeffrey E. Janso,[§] Louis R. Barrows,[‡] and Eric W. Schmidt*,[†]

[†]Department of Medicinal Chemistry and [‡]Department of Pharmacology and Toxicology, University of Utah, Salt Lake City, Utah 84112, United States

[§]Natural Products, Worldwide Medicinal Chemistry, Pfizer Worldwide Research and Development, Groton, Connecticut 06355, United States

Supporting Information

ABSTRACT: Strategies are needed for the robust production of cryptic, silenced, or engineered secondary metabolites in fungi. The filamentous fungus *Fusarium heterosporum* natively synthesizes the polyketide equisetin at >2 g L^{-1} in a controllable manner. We hypothesized that this production level was achieved by regulatory elements in the equisetin pathway, leading to the prediction that the same regulatory elements would be useful in producing other secondary metabolites. This was tested by using the native *eqxS* promoter and *eqxR* regulator in *F. heterosporum*, synthesizing heterologous natural products in yields of ∼1 g L^{-1}. As proof of concept for the practical application, we resurrected an extinct pathway from an endophytic fungus with an initial yield of >800 mg L^{-1}, leading to the practical synthesis of a selective antituberculosis agent. Finally, the method enabled new insights into the function of polyketide synthases in filamentous fungi. These results demonstrate a strategy for optimally employing native regulators for the robust synthesis of secondary metabolites.



Because of new technologies in sequencing and bioinformatics, it is now relatively trivial to identify novel biosynthetic pathways to secondary metabolites in cultivated organisms and in the environment.[1] Research in this area has led to a renewed appreciation that many secondary metabolites have yet to be described and that a vast resource awaits discovery.[2,3] Although pathway identification is simple, discovering the compounds produced by these pathways remains much more challenging, especially when the pathways are silent or encoded in the genomes of uncultivated organisms. To obtain the new compound, in many cases, the identified genes must be transferred to a new host and successfully expressed.[4] Despite many advances in technology, this is still nontrivial.

To date, many heterologous expression hosts have been developed for secondary metabolite production.[4−8] Most of these focus on expression of bacterial biosynthetic pathways, for which numerous hosts exist. For eukaryotic pathways, such as those from filamentous fungi, genes have been heterologously expressed in *Escherichia coli*, *Saccharomyces cerevisae*, *Aspergillus oryzae*, and several other yeasts and filamentous fungi.[5−7,9,10] Practical scale production of fungal compounds in bacteria has proven to be challenging.[4] *S. cerevisae* provides a robust platform but with relatively modest purified yields. Additionally, it does not handle introns well, requiring that introns be removed prior to expression. Similarly, heterologous expression platforms in various filamentous fungi have led to relatively modest yields, although in many cases fungal introns are tolerated, allowing genomic DNA to be employed directly.[5,9] These expression systems generally use housekeeping or related primary metabolic promoters that induce robust transcription of the desired genes, indicating that perhaps seeking improvements in level of transcription alone is insufficient to provide high levels of secondary metabolites, and suggesting room for the development of other strategies.

Here, we sought to take advantage of high-titer production of equisetin in the filamentous fungus *Fusarium heterosporum* ATCC 74349.[11,12] *F. heterosporum* produces equisetin at ∼2 g L^{-1} on corn grit agar (CGA), yet production is undetectable (10 ng L^{-1} detection limit) in many other types of media. We thus hypothesized that if an exogenous biosynthetic gene was placed under the control of the equisetin biosynthetic regulon, the heterologous compound would be synthesized in a yield similar to that observed for equisetin in the wild-type strain. Moreover, the highly controlled regulation of the equisetin locus might enable the production of compounds that are natively toxic to *F. heterosporum*. This strategy bears similarity to that previously used in actinomycetes to synthesize

Received: August 8, 2014
Published: September 16, 2014

   dx.doi.org/10.1021/sb500296p | *ACS Synth. Biol.* 2015, 4, 625−633

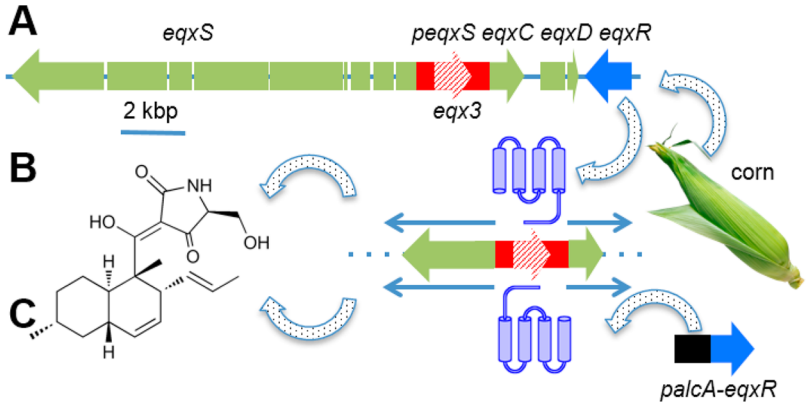


Figure 1. Model of the regulation of equisetin production. (A) The wild-type equisetin biosynthetic gene cluster, *eqx*, contains genes encoding biosynthetic proteins (green), a transcription factor (blue), and a promoter region used in this study (*peqxS*, red). (B) In the wild-type *F. heterosporum*, equisetin is produced when the fungus is grown on corn-derived media. A model is that the transcription factor gene *eqxR* is transcribed, leading to transcription of the pathway via *peqxS*. Shown is transcription of *eqxS* and *eqxC*, which together produce the desmethyl-equisetin analog, trichosetin, which is shown at left. (C) In support of this transcriptional model, when the heterologous promoter *palcA* is placed in front of *eqxR*, production of equisetin becomes constitutive and no longer depends upon growth on corn.

polyketides[13] but stands in contrast to a commonly used strategy in filamentous fungi, wherein promoters or regulators from primary metabolism are used to produce recombinant polyketides.

Using the model described below, we produced several heterologous fungal polyketides in *F. heterosporum*. In fungi, polyketides are biosynthesized by iterative decarboxylative condensation of malonyl units.[14,15] The minimal domains that make up the polyketide synthase (PKS) include acyltransferase (AT) domain, which selects the substrate; acyl carrier protein (ACP), which tethers the growing chain during extension; and the ketosynthase (KS) domain, which catalyzes the condensation reaction.[14−16] In addition, PKS enzymes may contain other modifying domains such as the methyltransferase (MT) domain, which introduces varying patterns of C-methylations along the polyketide backbone. For the reducing-type PKS, varying levels of reduction are achieved by ketoreductase (KR), dehydratase (DH), and enoyl reductase (ER) domains.[14] In general, all domains required to produce a fungal polyketide are contained on a single polypeptide, but for several PKSs, the ER is trans-acting and is translated as a separate polypeptide from the other PKS domains.[17−20] In the latter case, the trans-ER protein is required to produce the natural product. In many fungi, PKS modules exist as hybrids with nonribosomal peptide synthetases (NRPS), resulting in formation of polyketides fused to amino acids.[10,15,21,22] Unlike standalone PKSs that rely on hydrolysis via a thioesterase domain (TE) for product release, several PKS-NRPSs use a terminal reductase (R) domain, which catalyzes a Dieckmann reaction in tandem with product release to form a tetramic acid ring.[23,24]

Here, we describe a platform that enables production of heterologous fungal secondary metabolites in high titers (~1 g L^{-1} unoptimized yield). The platform was applied to several problems in biosynthesis and drug discovery, including resurrection of a silenced pathway of potential use in tuberculosis.

RESULTS AND DISCUSSION

Design of Expression Strategy. The strategy was based upon prior knowledge of the regulation of equisetin production. Equisetin biosynthesis requires the coordinated action of the PKS EqxS, the auxiliary ER EqxC, and the N-MT EqxD (Figure 1).[11] The first two of these are synthesized from genes that are divergently transcribed from a promoter region, *peqxS*. A regulatory transcription factor, EqxR, drives production of equisetin.[11] On CGA media, a large amount of equisetin is produced, while on other media such as potato dextrose agar (PDA; Difco) or potato dextrose broth (PDB; Difco), equisetin is not produced.[11] Equisetin is constitutively produced on CGA and slowly accumulates over a 21-day period. We thus selected *peqxS* as a platform for producing metabolites, under control of EqxR. In this conception, the *peqxS* region also contains a hypothetical gene, *eqx3*, which is commonly found in equisetin-like pathways. For simplicity, this region was left in place in the current study. In some vectors, we used only one-half of *peqxS* to express genes in one direction; in this series, *eqx3* is disrupted, and the promoter region is referred to as *peqxS′*. In others, the bidirectional promoter region *peqxS* was used; in these cases *eqx3* is not disrupted.

While the wild-type regulator *eqxR* was well controlled to produce equisetin solely on CGA, we envisioned applications in which we might desire more rapid production in liquid media. In previous work, we fused the *alcA* promoter (*palcA*) with *eqxR* (Figure 1C).[11] While *palcA* is inducible, it exhibits significant basal expression. Indeed, *palcA-eqxR* led to constitutive production of equisetin even in broth.[11] However, production yields were lower in broth, and there was a concern that toxic metabolites may delay or halt fungal growth under control of *palcA-eqxR*. By contrast, an advantage of the PDB method was that production could be achieved in 5−7 days, rather than the 21+ days used with CGA. The strategy thus took advantage of this strong control of equisetin production. If toxicity was a concern or if high yields were required, the wild-type *eqxR* construct was used. For a faster assessment of whether a recombinant metabolite could be produced, the *palcA-eqxR* construct could be employed. In this study, we employed two strategies to introduce *palcA-eqxR*: (1) we recently reported a modified *F. heterosporum* strain, Palc:eqxR, which overexpresses the equisetin positive regulator *eqxR* under control of the *alcA* promoter to allow equisetin production in broth culture after 5 days, whereas normally no equisetin is produced in liquid broth;[11] (2) we included the *palcA-eqxR*

dx.doi.org/10.1021/sb500296p | *ACS Synth. Biol.* 2015, 4, 625−633

gene in an expression vector that could be transferred to *F. heterosporum*, along with desired secondary metabolic genes.

We had two major concerns in terms of practical genetics: could we repetitively use a single promoter element, and could we process different classes of introns? *F. heterosporum* largely undergoes ectopic recombination in our hands, and it is quite difficult to obtain homologous recombinants. For this reason, we believed that repetitive use of *peqxS* would be tolerated by the strain, enabling production of more than two gene products without recombination. Indeed, in this study, we show that multiple copies of this element are stable in recombination, meaning that the strategy is scalable to recapitulate multigene biosynthetic pathways. We were also concerned about the ability to splice different types of introns. While this will presumably always be a problem in some cases, here, we show that a quite divergent set of introns can be processed. However, our overall strategy mainly uses artificially spliced, intron-free DNA.

Our construction strategy involved using shuttle vectors, in which we could employ yeast recombination in *S. cerevisiae*[25] to build the desired vectors, *E. coli* to amplify those vectors, and *F. heterosporum* for production of compounds. Each vector thus required selection and replication elements for each strain (see Methods). We constructed and tested multiple types of selection markers so that multiple different vectors could be inserted into *F. heterosporum*. Previously, we used hygromycin (*hph*) and phleomycin (*ble*) resistance effectively.[11] Here, we also employed uracil auxotrophy as an additional selection marker.[26] By combining these elements, it becomes possible to use stepwise engineering to insert multiple copies of genes under control of *peqxS*. In addition, it is always possible to insert multiple copies of the promoter into a single vector prior to transformation into *F. heterosporum*, making the strategy highly scalable.

Finally, we used an 8-cutter restriction endonuclease site, either AscI or PacI, to linearize the vectors before fungal transformation. These sites were selected so that the linearized vectors would contain the complete promoter-heterologous gene construct in the correct order. Otherwise, there was a danger of integration with a disrupted reading frame. In addition, we put these sites in some cases between selectable markers and synthetic green fluorescent protein (sGFP) to ensure that the vector was inserted in the correct manner. However, in the event, we never found a rearrangement that necessitated using sGFP. In *F. heterosporum*, linearized vectors integrated ectopically and intact.

Below, we describe a stepwise application of these principles to construct a heterologous expression platform. (1) The basics of transcription using both wild-type *eqxR* and *palcA-eqxR* were examined using sGFP. Although simple chemical analysis of equisetin production previously revealed fundamental aspects of pathway regulation, employing sGFP enabled a direct translational readout that would complement our understanding of *eqxR* and *peqxS*. (2) We desired to test this strategy by adding a single, discrete, heterologous PKS that leads to a known and well characterized product. We selected the *cpaS* gene from the *Aspergillus flavus* cyclopiazonic acid pathway for this purpose. Not only is *cpaS* exceptionally well characterized,[9,27] but *A. flavus* is in a different class (Eurotiomycetes) from *Fusarium* (Sordariomycetes). (3) We next tested the applicability to systems requiring two genes for production of compounds. In the event, we used both the homologous equisetin pathway[11] and the heterologous lovastatin pathway, again from an *Aspergillus*.[19] By examining equisetin production, we hoped to determine whether chromosomal location impacted production; in the event, it did not. The lovastatin PKS genes are well characterized[17,19,28] and biomedically of great importance, making this an interesting target for production. Moreover, only two proteins are required to synthesize the complex core of the molecule. (4) We aimed to produce more than two genes and to test the practical application of the platform to a real-world problem. This involved production of the "extinct" metabolite, pyrrolocin. Taken together, these approaches fully define the application of equisetin regulatory elements to produce diverse secondary metabolites in high yield.

Knockout of the *eqx* Locus. We desired to compare expression in the wild-type *F. heterosporum* with expression in an *eqx* knockout strain. Although *eqxC* has been previously deleted, *eqxS* was difficult to delete. Here, we created a knockout vector in which *eqxS*, *eqxC*, *eqxD*, and the promoter region *peqxS* were deleted (Figure 2). The resulting strain,

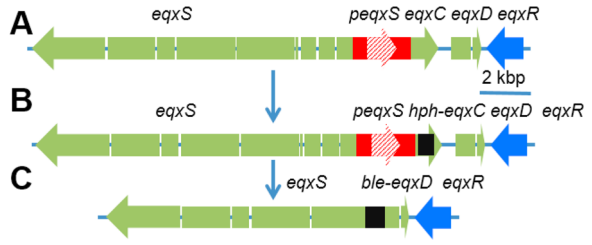


Figure 2. Knockout of *eqx* genes. (A) The wild-type gene cluster. (B) Previously, *eqxC* was deleted with *hph*. (C) Here, this *eqxC* knockout was extended to delete *peqxS* and *eqxC* and to disrupt the 5′-regions of *eqxS* and *eqxD*, using the *ble* marker. This knockout required the prior deletion of *eqxC* and did not work in the wild-type strain.

FusΔeqxS was confirmed to be an *eqx* cluster knockout by PCR (Supporting Information Figure S4). Subsequent genome sequencing of FusΔeqxS revealed that the *eqxS* gene was knocked out in the 5′ end and that the deletion vector was inserted into a total of 3 locations in the genome (Supporting Information Figure S5).

***eqxS* Promoter Drives Expression of sGFP in Recombinant Strain.** To test whether the *peqxS* promoter sequence could be used to express exogenous genes, and to determine timing and control of heterologous expression, the *peqxS′* was fused to sGFP and transferred to wild-type *F. heterosporum* in vector FH-1 (Figure 3A; Supporting Information Figure S1). After 21 days on CGA, sGFP was observed in fungal filaments when examined by confocal fluorescence microscopy. Moreover, the fungus was fluorescent over the full expression period on CGA, but it lacked any observable fluorescence on PDA. This was a clear indication that *peqxS′* was sufficient for heterologous expression. It also further supported previous observations that production was constitutive on CGA but completely shut off on PDA and other media. By contrast, when sGFP was added to strains containing the leaky *palcA-eqxR* gene (Figure 3B), the resulting *F. heterosporum* strain was constitutively fluorescent on all media.

Expression of the *cpaS* Gene from *Aspergillus flavus* Produces Expected Metabolite. To determine whether the *peqxS′* could lead to biosynthesis of new compounds, the well-characterized *cpaS* gene from *Aspergillus flavus* was cloned into FH-1 to make pHygB-Cpas. The resulting *F. heterosporum*

dx.doi.org/10.1021/sb500296p | *ACS Synth. Biol.* 2015, 4, 625−633

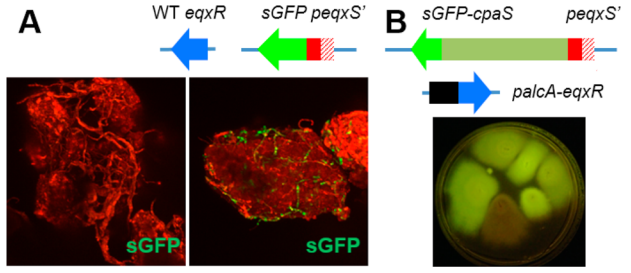


Figure 3. Monitoring *peqxS'*-driven transcription with sGFP. (A) *sGFP* was placed immediately downstream of *peqxS'* and cloned into the wild-type fungus. The wild-type control (left) was compared with the transformed fungus (right) on CGA. Shown is a single corn grit from a 21-day CGA culture, by confocal microscopy. Fluorescence was constitutive on CGA, but not visible on other media. Fungal filaments can be seen glowing green, over the red autofluorescent background. (B) *sGFP* was tethered to the gene, *cpaS*, downstream of the *peqxS'* promoter. The resulting vector also contained the constitutive *palcA-exqR* regulator. Shown here is a photograph of constitutively fluorescent fungal colonies grown on PDA, visualized on a Dark Reader.

mutant, Peqx:Cpas, was then cultured in potato dextrose broth (PDB) for 7 days. The predicted product, cAATrp **1** was identified by high pressure liquid chromatography (HPLC) (Figure 4A) and isolated. Comparison of the ^{1}H NMR spectrum and the molecular formula of **1** with those of the previously reported cAATrp showed that they were identical (Supporting Information Figure S13).[23,27] This confirmed the value of *peqxS'* in producing heterologous compounds.

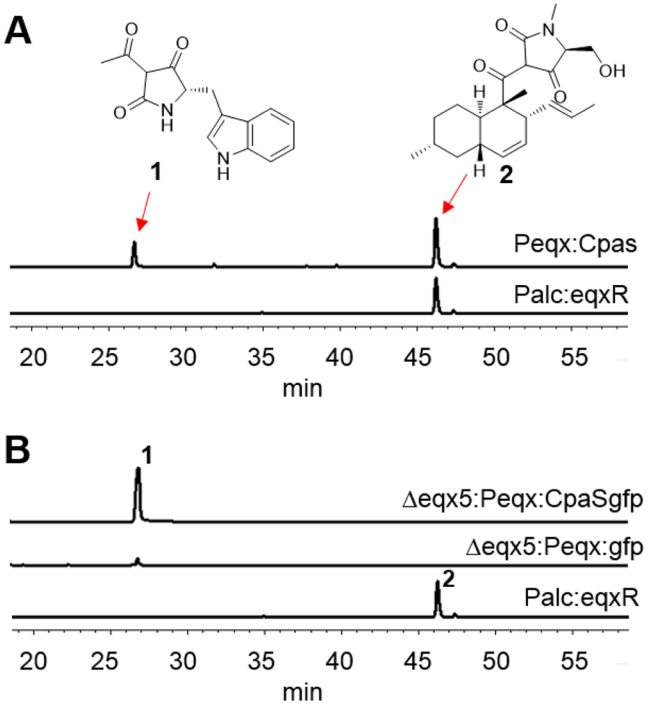


Figure 4. Expression of CpaS and production of cAATrp (**1**). (A) Expression in the wild-type *F. heterosporum* strain. HPLC-diode array detection (DAD) analysis of crude extracts of PDB cultures of Peqx:Cpas mutant and the Palc:eqxR control. The eqxS promoter drives expression to avail expected product **1**. Coproduction of equisetin **2** is observed. (B) Expression in *eqx* knockout strain. Production of **1** is improved, and equisetin **2** is no longer observed.

However, high levels of equisetin **2** were coproduced (Figure 4A).

In other systems, it has been shown that eliminating the production of undesired metabolites increases substrate flux to the target pathway resulting in higher yields of the desired compounds.[5] In our case, we reasoned that deleting the equisetin biosynthetic genes would not only increase flux of the building block, malonyl CoA, to polyketide-type heterologous pathways, but also ease downstream target compound purification and analysis. The *A. flavus cpaS* gene was cloned into vector FH-2 (Supporting Information Figure S2), which contained a copy of *palcA-eqxR* for constitutive expression in broth. An sGFP tag was fused to the C-terminus of *cpaS*. After transformation of this construct into the *eqx* knockout strain FusΔeqx5, we isolated a mutant Δeqx5:Peqx:CpaSgfp that glowed bright yellow on a dark reader (Figure 3B). This strain was cultured in 100 mL PDB for 7 days, extracted, and the extract analyzed by HPLC-DAD to show that **1** was robustly produced, while equisetin was absent (Figure 4B). Production of **1** was measured to be 100 mg L^{-1} by HPLC-DAD in comparison with a standard curve. This is an order of magnitude greater than the previously reported production of the same compound when the *cpaS* was expressed in *A. oryzae* under control of a housekeeping promoter.[27] In addition, this experiment provided some support for our hypothesis, since the yield in the deletion mutant was increased by approximately 4-fold in comparison to production in the wild-type equisetin producer.

When the Δeqx5:Peqx:CpaSgfp mutant was cultured for 21 days on CGA, production of **1** rose to over 1 g kg^{-1} without any optimization of production conditions. *A. flavus* is quite phylogenetically distant to *F. heterosporum*, with the former in Class Eurotiomycetes and the latter in Class Sordariomycetes, indicating that the strategy may be widely applicable to fungal metabolites from different groups.

Use of the Divergent Promoter for Simultaneous Dual Gene Introduction. In initial experiments, only one side of *peqxS* was used, comprising 1.5 kbp of gene sequence. To express two genes, *peqxS* was synthesized, containing the entire ∼2.5 kbp of the divergent promoter, to generate vector FH-3 (Figure 5A). Previously, we showed by knockout mutagenesis that *eqxC* was critical for the production of equisetin **2** and trichosetin **3**, and we proposed that it was the trans-acting ER for the equisetin pathway.[11] Here, we showed the direct involvement of *eqxS* in the biosynthesis of trichosetin **3**. Both *eqxS* and *eqxC* were cloned into vector FH-3 under control of the divergent *eqx* promoter. In initial experiments, *eqxS* was fused with a C-terminal sGFP tag so that we could readily confirm protein expression. The resulting vector, hpheqxC +eqxSgfp, was transformed into FusΔeqx5. The isolated transformants were brightly fluorescent, demonstrating appropriate gene expression under control of the divergent promoter. In addition, the fluorescence was constitutively obtained on PDA, further demonstrating the constitutive regulation of the pathway under control of leaky *palcA*.

Dieckmann Cyclase Depends upon Unmodified C-Terminal R Domain. To our surprise, trichosetin was not detected by HPLC-DAD in the crude extracts of the isolated mutants (Figure 5B). Instead, close inspection of the LC/MS trace showed a new product **4** heavier than trichosetin by 18 Da, but with a similar fragmentation pattern (Figure 5C; Supporting Information Figure S11). Compound **4** was characterized by NMR experiments, including ^{1}H, ^{13}C,

dx.doi.org/10.1021/sb500296p | ACS Synth. Biol. 2015, 4, 625−633

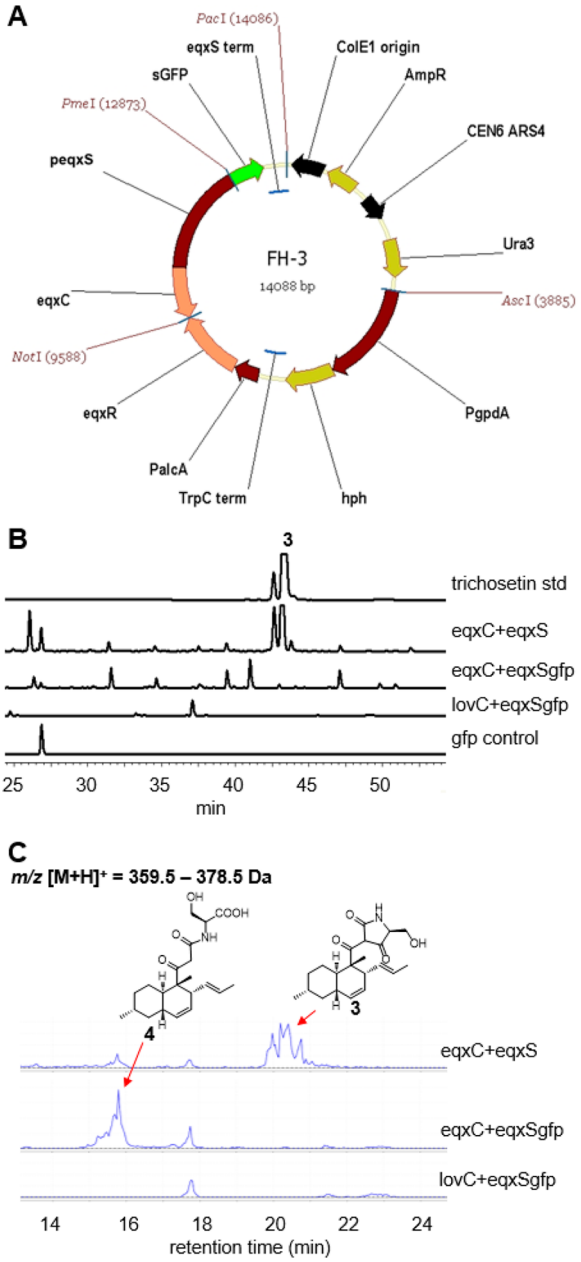


Figure 5. Trichosetin synthesis requires EqxC and unmodified EqxS. (A) Expression vector FH-3 designed with complete intergenic sequence *peqxS*, for dual expression of genes. Also shown are elements that permit cloning by recombination in *S. cerevisiae* and selection in *E. coli*. (B) Analytical HPLC of crude extracts of PDB cultures of FusΔeqxS transformed with *eqxSgfp* together with either *eqxC* or noncognate *trans*-ER *lovC*. Also shown is nontagged *eqxS* coexpressed with *eqxC* and the trichosetin standard. Trichosetin is only produced in the presence of *eqxC* and unmodified *eqxS*. (C) LC/MS analysis of crude extracts shows that a gfp tag on the C-terminus of EqxS interferes with formation of trichosetin and instead results in production of only the ring-open form **4**.

gCOSY, gHSQC, and gHMBC, to be the ring-opened derivative of trichosetin that had not undergone the terminal Dieckmann reaction.[23,24] This result implied that *eqxS* is indeed the equisetin synthetase, but the gfp tag interferes with the proper functioning of the reductase domain, preventing Dieckmann cyclization and instead promoting water-mediated hydrolysis of the intermediate. This was, in fact, found to be the case because, when *eqxS* was cloned without a C-terminal tag, trichosetin was robustly synthesized at wild-type levels (Figure 5B).

Production of Multiple Gene Products Leads to Synthesis of Lovastatin Precursors in *F. heterosporum*. To establish the utility of this new expression strategy for coexpression of heterologous genes, we cloned the well-studied lovastatin nonaketide synthase (*lovB*) from *Aspergillus terreus* together with its cognate *trans*-ER (*lovC*)[19] to make the hphlovC+lovBgfp plasmid. We also cloned only *lovB* into FH-3 to make the plasmid hpheqxC+lovBgfp. Transformation of these constructs independently into FusΔeqxS resulted in production of the expected products. Without its cognate *trans*-ER, the expressed LovB synthesized the previously reported truncated intermediates, **5** and **6** (Figure 6A).[19] Coexpression

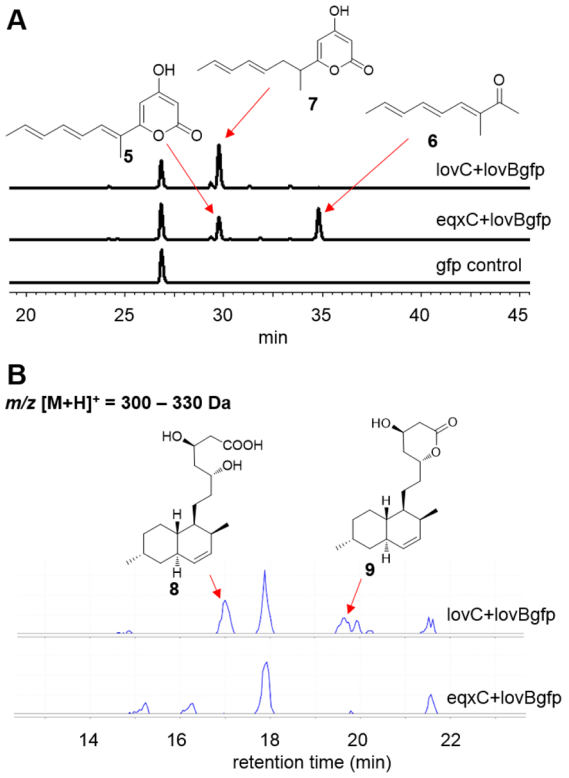


Figure 6. Dual expression with *peqxS* promoter reconstitutes pathway to lovastatin precursor. (A) Analytical HPLC of crude extracts of PDB cultures of FusΔeqxS transformed with *lovBgfp* together with *lovC* or a noncognate trans-ER *eqxC*. In the presence of *eqxC*, *lovB* produces the polyene pyrone **5** and ketone **6**; and more reduced pyrone **7** with *lovC* coexpression. (B) LC/MS analysis of crude extracts shows formation of the lovastatin precursor, dihydromonacolin L acid **8** and the lactone **9** when *lovBgfp* is coexpressed with *lovC*. EqxC is not able to complement LovB to form **8** or **9**.

of *lovB* with *lovC* produced the expected reduced metabolites **7** (Figure 6A), the dihydromonacolin L acid **8**, and the lovastatin precursor lactone **9** (Figure 6B).[19,29] These metabolites were characterized by comparing their liquid chromatography/mass spectroscopy (LC/MS), ultraviolet (UV), and ^{1}H NMR data to previous reports (Supporting Information Figures S8–S10, S12–S14). This confirmed that, indeed, the intergenic sequence between *eqxS* and *eqxC* could guide transcription divergently to heterologously coexpress two genes. The purified yield of **9** was 130 mg kg^{-1}, and **8** was produced in about equal

dx.doi.org/10.1021/sb500296p | ACS Synth. Biol. 2015, 4, 625–633

amounts, indicating that the initial unoptimized yield should exceed ~300 mg kg^{-1}.

EqxS Requires Cognate *trans*-ER. In the course of cloning and analyzing *lovC*, we coexpressed *lovC* and *eqxS* under control of the divergent promoter. No trichosetin was produced under these conditions (Figure 5B). LovC could not complement EqxC, indicating that EqxS could not interact with the noncognate ER. Further, the complementary experiment in which EqxC was coexpressed with LovB failed to yield reduced intermediates (Figure 6A). By contrast, the ERs were fully functional when coexpressed with their cognate PKS proteins.

Resurrection of Antituberculosis Agent from a Silenced Biosynthetic Pathway. In collaboration with scientists at Wyeth (now Pfizer), researchers in the Barrows lab cultivated an endophytic fungus, designated strain NRRL 50135, obtained in Papua New Guinea as part of their International Cooperative Drug Discovery Group (ICBG) project. The crude extract was potently active against *Mycobacterium tuberculosis*, and assay-guided purification led to the identification of a novel compound, pyrrolocin A (**10**), as the active principle. In 2007, a tentative structure **10** was assigned based upon HRESIMS and NMR (Figure 7A), but the amount of compound was not sufficient to assign the stereochemistry or to perform further activity tests. Unfortunately, the fungus ceased producing the compound after the initial isolation experiments, which is an extremely common phenomenon in natural products research, so that the project could not be further pursued or published. Numerous attempts at modifying production conditions, using many different published methods, failed to resurrect the synthesis of this potentially important compound. This presented the perfect challenge to test the application of our new production platform. The goal was to resurrect production in *F. heterosporum*, in tandem with greatly increasing the production level so that the compound could be further developed.

We sequenced the NRRL 50135 genome, which after assembly was found to be 54.1 Mbp on 5809 contigs (calculated GC content is 47.5%). Autoannotation software predicted 17722 proteins. From the preliminary structure of **10**, we predicted its biosynthesis would be similar to that of equisetin,[11] requiring a PKS-NRPS, *trans*-ER, and MT. BLAST analysis of the predicted proteins using EqxS as a query produced only 2 PKS-NRPS genes in the entire genome. When the database was queried with EqxC, the top hit was on the same 50.7 kbp contig as the identified top hit obtained with the EqxS query (Figure 7B). These genes were subsequently designated *prlS* and *prlC*, in analogy with the equisetin nomenclature. However, we could not find an MT that coclustered with either of the PKS-NRPS genes. Instead, distantly similar genes (37% identity to the equisetin *N*-methyltransferase, EqxD) were found unclustered with any PKS.

Since production of **10** was predicted to require coexpression of more than two genes, we created a derivative strain of FusΔeqx5, in which the strain was transformed into the uracil auxotroph, FusΔeqx5ΔpyrG10, by directed knockout of the orotidine 5′-phosphate decarboxylase gene (*pyrG*). A complement vector for this auxotroph was constructed using the native *F. heterosporum pyrG*, to create expression vector FH-4 (Supporting Information Figure S3). The *prlS* and *prlC* sequences were then cloned into FH-4 and used to transform strain Δeqx5ΔpyrG10 to prototrophy. An isolated mutant ΔpyrG:Peqx:prlS+prlC was cultured in 250 mL PDB for 5 days. The crude extract contained pyrrolocins B (**11**) and C (**12**), the desmethyl derivatives of **10** (444 Da $[M + H]^+$).[33] Compounds **11** and **12** (Figure 7C) were produced in a 2:1 ratio, totaling >800 mg kg^{-1} on CGA, which is greatly in excess of what was initially found for **10** in the native producer before production was lost (~50 mg L^{-1}).

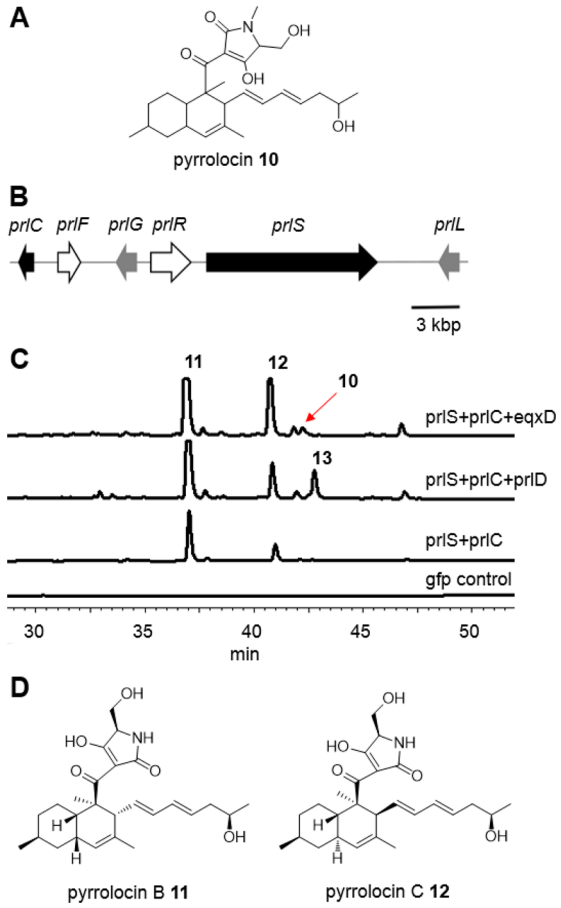


Figure 7. (A) Preliminary structure of antituberculosis agent **10** with mass 458 Da $(M+H)^+$ initially isolated from endophytic fungus NRRL 50135. (B) The identified candidate gene cluster for the biosynthesis of **10** after genome sequencing contains a PKS-NRPS hybrid gene *prlS*, an enoyl reductase *prlC*, two transcription factors *prlF* and *prlR*, and two exporter genes *prlG* and *prL*. (C) Analytical HPLC of crude extract of FusΔeqx5ΔpyrG10 transformed with both *prlS* and *prlC* driven by the *eqxS* promoter shows synthesis of two new products **11** and **12** with corresponding mass of 444 Da $(M+H)^+$. Further introduction of methyltransferases *prlD* and *eqxD* led to synthesis of new minor products **13** and **10**, respectively. (D) Compounds **11** and **12** were found to have antituberculosis activity and to have the core structure of **10**.[33]

Two MTs were cloned in attempts to methylate the tetramate ring. The first, PrlD, was the top hit resulting from BLAST analysis of NRRL 50135 when EqxD was used as the query (36% identity). *prlD* was from a contig that was not linked with *prl*, meaning that it might result from a separate biosynthetic pathway. EqxD itself was also coexpressed with PrlS and PrlC. Each MT was cloned into complementary vector FH-1 and separately transformed into *F. heterosporum* ΔpyrG:Peqx:prlS+prlC. The mutants were cultivated for 7 days. Analysis of the crude extracts by HPLC-DAD and LC/MS showed that both PrlD and EqxD led to synthesis of methylated products. Interestingly, the EqxD-expressing strain

dx.doi.org/10.1021/sb500296p | *ACS Synth. Biol.* 2015, 4, 625−633

produced a minor amount of authentic pyrrolocin A **10** (<10 mg kg^{-1}) in a background of ~800 mg kg^{-1} of nonmethylated **11** and **12**. The methylated derivative **13** produced by the PrlD-expressing strain did not match the NMR data for **10** and is therefore not the correct MT. No obvious MT exists in the sequenced genome that would be predicted to perform that transformation.

Conclusion. Heterologous expression is potentially the most universal solution to produce natural products from cryptic pathways and unculturable organisms.[4,5] We set out to develop such a production platform based on the equisetin biosynthetic regulon, which would enable production of important metabolites with sufficient yield to ease purification and downstream assays. This platform was designed based upon the hypothesis that high-level transcription alone was insufficient to produce high levels of secondary metabolites. Instead, exploitation of a natively high-producing biosynthetic pathway would provide the ideal environment for secondary metabolite biosynthesis. In developing and testing this platform, we designed derivative strains of *F. heterosporum* and a suite of expression vectors into which various heterologous genes could be cloned. These new tools were validated by producing a range of fungal polyketides from widely different fungi, in unoptimized yields ranging from several hundred mg to >1 g per liter. In fact, we found that cultures on just 50 g of CGA provided enough material for complete characterization of novel compounds. We provided proof of concept by resurrecting an important silenced pathway.

Another advantage of this platform is that it seems to process introns from a broad phylogenetic diversity of fungi. *prlS* and *prlC* contained introns that were properly processed (otherwise interrupting the reading frame of the proteins), demonstrating the breadth of the platform in handling introns from diverse strains. As has been found in other heterologous expression strategies, a disadvantage is that pathway shunt products are often produced along with the major natural product. However, in this case, the abundant amount of desired products compensated for this problem.

We have demonstrated practical application of this platform in reviving antituberculosis activity previously observed from pyrrolocin A, a product of the NRRL 50135 strain. Direct cloning of *prlS* and *prlC* from genomic DNA of strain NRRL 50135 into our expression vectors and subsequent high-level production of related active compounds **11** and **12** demonstrated the ease with which such eukaryotic genes ridden with introns could be characterized. This work enabled the structures of the pyrrolocins to be fully elucidated and the biological activity to be characterized.[33]

This strategy is widely applicable to many types of natural product synthesis in different organisms. The key factors that led us to select *F. heterosporum* as the production host were as follows: (1) the robust synthesis of the native compound, which meant that even if the recombinant yield were greatly reduced it should still be sufficient for chemical analysis; (2) the strict controllability of compound production, where essentially no natural product is produced on most media types, leading to a reduction in problems related to toxicity; (3) the apparent constitutive expression of the natural pathway, where compounds are slowly produced, exported, and accumulated in the medium over the course of weeks on CGA. It is likely that other systems with similar features would be amenable to the same approach.

A few biochemical observations were also enabled by these studies. During the complementation of the *eqx* knockout strain with *eqxC* and *eqxS*, we directly show that both these genes are required for equisetin production. However, we found that R domain function is altered when fused to an sGFP tag. Instead of the expected tetramic acid **3**, we obtained the ring-open **4**. We speculate that sGFP alters the structure of the R domain to prevent either transfer from the T domain or to allow water to enter the R domain active site. The tetramic acid derivative is nearly instantaneously formed from thioesters under neutral buffer conditions in water,[24] making it especially remarkable that the linear form can be obtained from thiotemplated synthesis. Potentially, if desired, this problem might be circumvented by experimenting with different types of linkers that do not disrupt the R domain. A cleavable linker strategy may also be feasible. However, in our hands, we found that sGFP is optional; for example, the complete vector is usually integrated intact in *F. heterosporum*. We also found that the ER proteins cannot be crossed between these pathways but that the wild-type ERs are required at least in these cases. The interaction of these ER proteins is also of interest, since having the correct set of protein partners is essential in the synthesis of the desired natural products.

In the wild-type fungus NRRL 50135, **10** was the major compound produced, with a small amount (<1% estimated from HPLC-MS) of **12** as a side-product.[33] Based upon the sequenced gene cluster and the recombinant expression performed here, it is clear that the combination of PKS-NRPS and auxiliary ER lead to the formation of **12**. Unlike our findings here, the wild-type fungus did not produce any *cis*-decalin product, such as **11**. It is possible that one of the other hypothetical genes in the gene cluster could be responsible for this discrepancy and may act as the pyrrolocin Diels–Alderase, but there are also other interesting possibilities. Of note, no obvious MT was present in the *prl* cluster that might produce **10**. The closest homologue of EqxD from the genome did not produce **10**. By contrast, EqxD itself produced ~10 mg kg^{-1} of **10**. This minor production, using a combination of heterologous and homologous proteins, would likely not have been observable if starting with a less efficient expression system. This emphasizes the value of starting with a high-yielding platform. The yield was still sufficient for biological and chemical characterization from a single 1 kg scale experiment. The reaction with EqxD is also remarkable in that **12** contains D-Ser, while **3** contains L-Ser, and the decalin ring is also enantiomeric between **3** and **12**. This indicates that, perhaps, EqxD has fairly relaxed substrate selectivity.

METHODS

Cloning of Vectors and Expression Plasmids. Standard PCR techniques were employed and plasmid construction carried out as previously described.[11] Details of vector construction can be found in the Supporting Information. Vector images were generated with Vector NTI software (Invitrogen).

Transformation of *Fusarium heterosporum* was done as previously described except that protoplasts were prepared from 8 to 10-h germinating spores.[11] When antibiotic selection was required, hygromycin or phleomycin was added to media at 150 μg mL^{-1}. For uracil auxotroph selection, protoplast regeneration agar lacking uracil was prepared containing 1 M sucrose, 0.02% yeast extract without amino acids, 0.02% BSM supplement, and 1% agar.

Fungal Mutagenesis. Knockout of *eqxS* FusΔeqx5 was constructed by transforming FusΔeqxC with knockout vector ClusterPhleoKO (Supporting Information) and transformants selected on phleomycin. The knockout cassette was made of the phleomycin resistance marker flanked by sequences homologous to regions within *eqxS* and *eqxD* (Figure 2). Isolated transformants were counter-screened for hygromycin sensitivity and verified by colony PCR. The genome of the identified knockout FusΔeqx5 was extracted and sequenced. The reads were aligned to the FusWT reference genome[11] with Novoalign and output visualized with the integrative genomics viewer[30] (Supporting Information Figure S5).

Generating the Uracil Auxotroph. The pyrG knockout cassette in the TOPO-pyrGKO plasmid was made by cloning a randomly selected sequence (first exon of *lovC*) into the *F. heterosporum pyrG* sequence. Translation of this sequence results in a truncated, nonfunctional PyrG. FusΔeqx5 was transformed with TOPO-pyrGKO, and the protoplasts were regenerated for 72 h at 30 °C before plating on selection medium made of Czapek Dox Broth, 5-FOA (4 g L^{-1}), uracil (1.12 g L^{-1}), uridine (140 mg L^{-1}), and 1.5% agar. Transformants were cross-streaked on Czapek−Dox agar with and without uracil to identify auxotrophs. Diagnostic PCR for homologous integration was done to confirm FusΔeqx5Δ-pyrG10 as a true knockout (Supporting Information Figure S6).

Fluorescence Microscopy. Visualization of gfp expressing mutants on CGA was done on an Olympus FV1000 spectral confocal microscope. Visualization of gfp expressing mutants on PDA was done on a Dark Reader (Clare Chemical).

Genome Sequencing and Analysis. All fungal genomes were extracted with the DNeasy Plant Minikit (Qiagen) and sequenced at the University of Utah Huntsman Cancer Institute sequencing facility on an Illumina HiSeq 2000.

The raw reads from NRRL 50135 genome sequencing were assembled using the SPAdes de novo genome assembler[31] with a kmer range from 31 to 85. Autoannotation of the genome was done using Augustus.[32] To locate the pyrrolocin biosynthetic cluster, BLAST analysis was done using EqxS, EqxC, and EqxD sequences as queries.

Chemical Analysis. Selected transformants were screened for compound production by culture in potato dextrose broth (PDB) for 7 d at 30 °C with shaking at 180 rpm. Spores of compound expressing mutants were then inoculated on to corn grit agar and incubated at room temperature for 21 d. Extraction of filtered PDB broth was done with ethyl acetate containing 1% acetic acid, and CGA was wholly extracted with acetone. Solvents were removed under vacuum, and the crude extracts analyzed by HPLC-DAD and LC/MS using C18 chromatography.

Purification of cAATrp 1. The crude extract from a PDB culture (250 mL) of Δeqx5:Peqx:CpaSgfp was purified by preparative HPLC (4 mL min^{-1}; 5−70% acetonitrile/water-0.05% TFA in 35 min). The pooled fractions were dried under vacuum to afford pure **1** (9 mg), which was then analyzed by ^{1}H NMR and LC/MS (Supporting Information Figures S7 and S13), in comparison to previously published reference data.[23,27] A standard curve was generated using pure **1** by HPLC-DAD and production was quantified from crude extracts of 100 mL PDB cultures and 50 g CGA cultures of Δeqx5:Peqx:Cpasgfp to average 100 mg L^{-1} and 1.25 g/kg (±0.24 g kg^{-1}, n = 3), respectively.

Purification of 4. CGA culture of Δeqx5:Peqx:eqxC +eqxSgfp (50 g) was extracted with acetone. The crude residue was then fractionated by flash chromatography on end-capped C18 with a methanol/water gradient and fractions screened by LC/MS. The fraction containing **4** was further purified by preparative HPLC to afford previously undescribed compound **4** (4.7 mg), which was characterized spectroscopically (see Supporting Information).

Purification of Dihydromonacolin L 9. The crude extract from 50 g CGA culture of Δeqx5:Peqx:lovC+lovBgfp was fractionated by flash chromatography on end-capped C18 with a methanol/water gradient. Fractions were screened by LC/MS and **9** was found to be contained in one fraction. This fraction was dried, and the residue was separated on silica column with 2:1 ethyl acetate/hexanes mobile phase. Further preparative HPLC purification afforded **9** as a white solid (6.6 mg). MS and ^{1}H NMR data were compared with those for the previously reported material.[19]

Purification of 10−13. To purify **11**, the crude extract from 200 g CGA culture was fractionated by flash chromatography on end-capped C18 using a methanol/water gradient. The fractions were analyzed by HPLC-DAD, and a portion (12.5%) of the fraction containing **11** was purified by several rounds of preparative HPLC to obtain pure **11** (8.6 mg; isolated yield 344 mg kg^{-1}). A similar strategy was used to obtain other derivatives.

■ ASSOCIATED CONTENT

Supporting Information

Plasmid construction; plasmids and primers used; NMR data; vector maps; PCR analyses; alignment of FusΔeqx5 illumina reads; LC/MS and UV data; ^{1}H and ^{13}C spectra. This material is available free of charge via the Internet at http://pubs.acs.org.

Accession Codes

The *prl* sequence has been deposited in GenBank, accession number KM107910.

■ AUTHOR INFORMATION

Corresponding Author

*Tel.: 801-585-5234; Email: ews1@utah.edu.

Notes

The authors declare no competing financial interest

■ ACKNOWLEDGMENTS

This work was funded by the National Science Foundation (NSF), 0957791. Collection and chemical analysis of the pyrrolocin-producing strain was funded by the National Institutes of Health (NIH), 5UO1TW006671-10.

■ REFERENCES

(1) Challis, G. L. (2008) Genome mining for novel natural product discovery. *J. Med. Chem. 51*, 2618−2628.

(2) Brakhage, A. A. (2013) Regulation of fungal secondary metabolism. *Nat. Rev. Microbiol. 11*, 21−32.

(3) Evans, B. S., Robinson, S. J., and Kelleher, N. L. (2011) Surveys of non-ribosomal peptide and polyketide assembly lines in fungi and prospects for their analysis *in vitro* and *in vivo*. *Fungal Genet. Biol. 48*, 49−61.

(4) Ongley, S. E., Bian, X., Neilan, B. A., and Muller, R. (2013) Recent advances in the heterologous expression of microbial natural product biosynthetic pathways. *Nat. Prod. Rep. 30*, 1121−1138.

(5) Chiang, Y.-M., Oakley, C. E., Ahuja, M., Entwistle, R., Schultz, A., Chang, S.-L., Sung, C. T., Wang, C. C. C., and Oakley, B. R. (2013) An

efficient system for heterologous expression of secondary metabolite genes in *Aspergillus nidulans*. *J. Am. Chem. Soc. 135*, 7720−7731.
(6) Kealey, J. T., Liu, L., Santi, D. V., Betlach, M. C., and Barr, P. J. (1998) Production of a polyketide natural product in nonpolyketide-producing prokaryotic and eukaryotic hosts. *Proc. Natl. Acad. Sci. U.S.A. 95*, 505−509.
(7) Pfeifer, B. A., Admiraal, S. J., Gramajo, H., Cane, D. E., and Khosla, C. (2001) Biosynthesis of complex polyketides in a metabolically engineered strain of *E. coli*. *Science 291*, 1790−1792.
(8) Li, C., Zhang, F., and Kelly, W. L. (2011) Heterologous production of thiostrepton A and biosynthetic engineering of thiostrepton analogs. *Mol. BioSyst. 7*, 82−90.
(9) Tokuoka, M., Seshime, Y., Fujii, I., Kitamoto, K., Takahashi, T., and Koyama, Y. (2008) Identification of a novel polyketide synthase−nonribosomal peptide synthetase (PKS−NRPS) gene required for the biosynthesis of cyclopiazonic acid in *Aspergillus oryzae*. *Fungal Genet. Biol. 45*, 1608−1615.
(10) Xu, W., Cai, X., Jung, M. E., and Tang, Y. (2010) Analysis of intact and dissected fungal polyketide synthase-nonribosomal peptide synthetase *in vitro* and in *Saccharomyces cerevisiae*. *J. Am. Chem. Soc. 132*, 13604−13607.
(11) Kakule, T. B., Sardar, D., Lin, Z., and Schmidt, E. W. (2013) Two related pyrrolidinedione synthetase loci in *Fusarium heterosporum* ATCC 74349 produce divergent metabolites. *ACS Chem. Biol. 8*, 1549−1557.
(12) Singh, S. B., Zink, D. L., Goetz, M. A., Dombrowski, A. W., Polishook, J. D., and Hazuda, D. J. (1998) Equisetin and a novel opposite stereochemical homolog phomasetin, two fungal metabolites as inhibitors of HIV-1 integrase. *Tetrahedron Lett. 39*, 2243−2246.
(13) Stevens, D. C., Hari, T. P. A., and Boddy, C. N. (2013) The role of transcription in heterologous expression of polyketides in bacterial hosts. *Nat. Prod. Rep. 30*, 1391−1411.
(14) Chooi, Y. H., and Tang, Y. (2012) Navigating the fungal polyketide chemical space: From genes to molecules. *J. Org. Chem. 77*, 9933−9953.
(15) Boettger, D., and Hertweck, C. (2013) Molecular diversity sculpted by fungal PKS-NRPS hybrids. *ChemBioChem 14*, 28−42.
(16) Walsh, C. T., and Fischbach, M. A. (2010) Natural Products Version 2.0: Connecting genes to molecules. *J. Am. Chem. Soc. 132*, 2469−2493.
(17) Ames, B. D., Chi, N., Bruegger, J., Smith, P., Xu, W., Ma, S., Wong, E., Wong, S., Xie, X., Li, J. W. H., Vederas, J. C., Tang, Y., and Tsai, S.-C. (2012) Crystal structure and biochemical studies of the trans-acting polyketide enoyl reductase LovC from lovastatin biosynthesis. *Proc. Natl. Acad. Sci. U.S.A. 109*, 11144−11149.
(18) Heneghan, M. N., Yakasai, A. A., Williams, K., Kadir, K. A., Wasil, Z., Bakeer, W., Fisch, K. M., Bailey, A. M., Simpson, T. J., Cox, R. J., and Lazarus, C. M. (2011) The programming role of trans-acting enoyl reductases during the biosynthesis of highly reduced fungal polyketides. *Chem. Sci. 2*, 972−980.
(19) Ma, S. M., Li, J. W. H., Choi, J. W., Zhou, H., Lee, K. K. M., Moorthie, V. A., Xie, X., Kealey, J. T., Da Silva, N. A., Vederas, J. C., and Tang, Y. (2009) Complete reconstitution of a highly reducing iterative polyketide synthase. *Science 326*, 589−592.
(20) Bergmann, S., Schuemann, J., Scherlach, K., Lange, C., Brakhage, A. A., and Hertweck, C. (2007) Genomics-driven discovery of PKS-NRPS hybrid metabolites from *Aspergillus nidulans*. *Nat. Chem. Biol. 3*, 213−217.
(21) Fisch, K. M. (2013) Biosynthesis of natural products by microbial iterative hybrid PKS-NRPS. *RSC Adv. 3*, 18228−18247.
(22) Maiya, S., Grundmann, A., Li, X., Li, S.-M., and Turner, G. (2007) Identification of a hybrid PKS/NRPS required for pseurotin A biosynthesis in the human pathogen *Aspergillus fumigatus*. *ChemBioChem 8*, 1736−1743.
(23) Liu, X., and Walsh, C. T. (2009) Cyclopiazonic acid biosynthesis in *Aspergillus* sp.: characterization of a reductase-like R* domain in cyclopiazonate synthetase that forms and releases cyclo-acetoacetyl-L-tryptophan. *Biochemistry 48*, 8746−8757.
(24) Sims, J. W., and Schmidt, E. W. (2008) Thioesterase-like role for fungal PKS-NRPS hybrid reductive domains. *J. Am. Chem. Soc. 130*, 11149−11155.
(25) Ma, H., Kunes, S., Schatz, P. J., and Botstein, D. (1987) Plasmid construction by homologous recombination in yeast. *Gene 58*, 201−216.
(26) Weld, R. J., Plummer, K. M., Carpenter, M. A., and Ridgway, H. J. (2006) Approaches to functional genomics in filamentous fungi. *Cell Res. 16*, 31−44.
(27) Seshime, Y., Juvvadi, P. R., Tokuoka, M., Koyama, Y., Kitamoto, K., Ebizuka, Y., and Fujii, I. (2009) Functional expression of the *Aspergillus flavus* PKS−NRPS hybrid CpaA involved in the biosynthesis of cyclopiazonic acid. *Bioorg. Med. Chem. Lett. 19*, 3288−3292.
(28) Auclair, K., Sutherland, A., Kennedy, J., Witter, D. J., Van den Heever, J. P., Hutchinson, C. R., and Vederas, J. C. (2000) Lovastatin nonaketide synthase catalyzes an intramolecular Diels−Alder reaction of a substrate analogue. *J. Am. Chem. Soc. 122*, 11519−11520.
(29) Xu, W., Chooi, Y. H., Choi, J. W., Li, S., Vederas, J. C., Da Silva, N. A., and Tang, Y. (2013) LovG: The thioesterase required for dihydromonacolin L release and lovastatin nonaketide synthase turnover in lovastatin biosynthesis. *Angew. Chem., Int. Ed. Engl. 52*, 6472−6475.
(30) Robinson, J. T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G., and Mesirov, J. P. (2011) Integrative genomics viewer. *Nat. Biotechnol. 29*, 24−26.
(31) Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., Lesin, V. M., Nikolenko, S. I., Pham, S., Prjibelski, A. D., Pyshkin, A. V., Sirotkin, A. V., Vyahhi, N., Tesler, G., Alekseyev, M. A., and Pevzner, P. A. (2012) SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol. 19*, 455−477.
(32) Stanke, M., Steinkamp, R., Waack, S., and Morgenstern, B. (2004) AUGUSTUS: A web server for gene finding in eukaryotes. *Nucleic Acids Res. 32*, W309−W312.
(33) Jadulco, R. C.; Koch, M.; Kakule, T. B.; Schmidt, E. W.; Orendt, A.; He, H.; Janso, J. E.; Carter, G. T.; Larson, E. C.; Pond, C.; Matainaho, T. K.; Barrows, L. R. Isolation of pyrrolocins A−C: *cis*- and *trans*-decalin tetramic acid antibiotics from an endophytic fungal-derived pathway. *J. Nat. Prod.* 2014, *Submitted*.