



Published in final edited form as:

ACS Chem Biol. 2018 November 16; 13(11): 3193–3205. doi:10.1021/acscchembio.8b00679.

Hybrid Transcription Factor Engineering Activates the Silent Secondary Metabolite Gene Cluster for (+)-Asperlin in *Aspergillus nidulans*

Michelle F. Grau[†], Ruth Entwistle[‡], Yi-Ming Chiang^{†,§}, Manmeet Ahuja^{†,∇}, C. Elizabeth Oakley[‡], Tomohiro Akashi^{||}, Clay C. C. Wang^{*,†,⊥}, Richard B. Todd^{*,#}, and Berl R. Oakley^{*,‡}

[†]Department of Pharmacology and Pharmaceutical Sciences, School of Pharmacy, University of Southern California, Los Angeles, California 90089, United States

[‡]Department of Molecular Biosciences, University of Kansas, Lawrence, Kansas 66045, United States

[§]Department of Pharmacy, Chia Nan University of Pharmacy and Science, Tainan City 71710, Taiwan

^{||}Division of OMICS analysis, Center for Neurological Diseases and Cancer, Nagoya University Graduate School of Medicine, Nagoya, Japan

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

[#]Department of Plant Pathology, Kansas State University, 4024 Throckmorton Plant Sciences Center, Manhattan, KS 66506, United States

Abstract

Fungi are a major source of valuable bioactive secondary metabolites (SMs). These compounds are synthesized by enzymes encoded by genes that are clustered in the genome. The vast majority of SM biosynthetic gene clusters are not expressed under normal growth conditions, and their products are unknown. Developing methods for activation of these silent gene clusters offers the potential for discovering many valuable new fungal SMs. While a number of useful approaches have been developed, they each have limitations and additional tools are needed. One approach, upregulation of SM gene cluster-specific transcription factors that are associated with many SM gene clusters, has worked extremely well in some cases, but it has failed more often than it has succeeded. Taking advantage of transcription factor domain modularity, we have developed a new approach. We have fused the DNA-binding domain of a transcription factor associated with a silent SM gene cluster with the activation domain of a robust SM transcription factor, AfoA.

*Corresponding Author clayw@usc.edu, rbtodd@ksu.edu, boakley@ku.edu.

[∇]Present Addresses Industrial Biotechnology Division, Reliance Technology Group, Reliance Industries Limited, Reliance Corporate Park, Thane Belapur Road, Ghansoli, Navi, Mumbai 400701, India.

Supporting Information

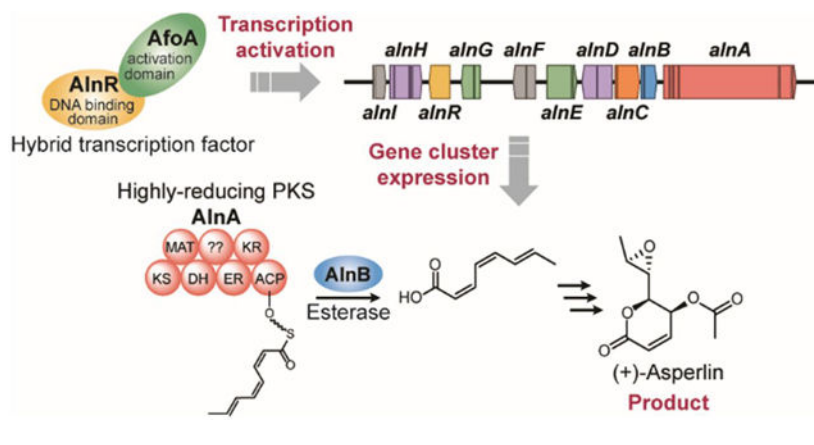
The Supporting Information is available free of charge on the ACS Publications website at DOI: [10.1021/acscchembio.8b00679](https://doi.org/10.1021/acscchembio.8b00679)

Detailed structural characterization including HIRESIMS and NMR spectroscopic data, hybrid transcription factor and reannotated gene sequences, and CAGE RNA-seq analysis for AN9221.

The authors declare no competing financial interest.

Expression of this hybrid transcription factor activated transcription of the genes in the target cluster and production of the antibiotic (+)-asperlin. Deletion of cluster genes confirmed that the cluster is responsible for (+)-asperlin production, and we designate it the *aln* cluster. Separately, co-induction of expression of two *aln* cluster genes revealed the pathway intermediate (2*Z*,4*Z*,6*E*)-octatrienoic acid, a compound with photoprotectant properties. Our findings demonstrate the potential of our novel synthetic hybrid transcription factor strategy to discover the products of other silent fungal SM gene clusters.

Graphical Abstract:



Secondary metabolites (SMs) are a rich source of medically important compounds. These compounds are not strictly required for viability, but they confer a selective advantage for the producing organisms, often by inhibiting important biological processes in their competitors in their native environments. These inhibitory activities are often medically useful. Some SMs have been used as medicines without modification (e.g. penicillin and lovastatin), other medically valuable compounds are derived from SMs, and yet additional compounds are inspired by SMs (e.g. with pharmacophores based on SMs).

Fungi have been an excellent source of medically valuable SMs¹ and, although many thousands of fungal SMs have been isolated, genome sequencing has revealed that the vast majority of fungal SMs have yet to be identified. Fungal genome projects have confirmed previous conclusions that the genes that encode the biosynthetic enzymes required for the production of particular compounds are clustered together in the genome and coordinately regulated. They have also revealed that for nearly all fungi examined the number of SM biosynthetic gene clusters greatly exceeds the number of compounds known to be produced by the fungus²⁻⁵.

The reason is that most SM gene clusters are silent under normal lab growth conditions, and developing strategies for stimulating expression of these clusters, such that their products can be isolated, has become an important research focus⁶. A number of strategies have been developed⁷⁻¹¹. These include alteration of the media on which the fungi are grown¹², deletion or overexpression of genes involved in chromatin packing^{13,14}, chemical perturbation of proteins that modify chromatin¹⁵⁻¹⁸, isolation of mutations in, or over expression of, genes such as *mcrA* and *laeA* that affect the production of many secondary

metabolites^{19,20}, co-incubation with other microorganisms²¹, and replacement of the promoters of SM biosynthetic gene cluster genes with regulatable promoters^{22,23}. All of these strategies have had notable successes, but none of them have proved to be universally effective. All of these strategies have been applied to *Aspergillus nidulans*, the fungus in which the secondary metabolome has been most extensively explored, resulting in the discovery of scores of compounds new to *A. nidulans* or new to science¹¹. Nevertheless, many, probably more than half, of the compounds produced by this organism have yet to be discovered.

One approach, that has been successful in some instances, takes advantage of the fact that many SM biosynthetic gene clusters contain a gene encoding a transcription factor that drives expression of all of the genes of the cluster. In some cases replacement of the promoter of the transcription factor with a regulatable promoter, and subsequent upregulation of expression of the transcription factor, results in strong expression of the genes of the cluster and high levels of production of the compound produced by the pathway^{24,25}. This approach can work spectacularly well, resulting in very high yields of secondary metabolites [e.g. > 2 g/L of asperbenzaldehyde²⁶], but the approach often fails. This approach was tried with 18 SM biosynthetic gene clusters in *A. nidulans*²². In three of these cases the approach was highly successful, but in the other cases no compounds, or unanalyzably small amounts of compound were produced.

There are several potential causes for failure of this approach. Some are intellectually trivial, but practically consequential. For example, the induction conditions may have been inadequate to stimulate the required level of production, or the transcription factor may have been incorrectly annotated such that the promoter was not driving expression, or was driving expression of only a portion of the transcription factor. These problems are compounded by the fact that transcription factors may work as heterodimers and if either of the two required transcription factors is inadequately expressed or incorrectly annotated, SM production will fail. These problems will undoubtedly be addressed by better annotation algorithms and data and by better expression systems. Other potential limitations of this approach are related to normal SM regulation. One theoretical, but logical, possibility is that interaction of the transcription factor with an unknown natural inducer might be required for strong activation. Another, well documented, factor is that the activity of transcription factors that drive expression of SMs can be regulated post-translationally and, when this is the case, simple upregulation is unlikely to result in high levels of expression of the genes of the regulated cluster. The best studied example of this is the AfIR transcription factor that drives expression of the sterigmatocystin gene cluster of *A. nidulans*. Regulation of AfIR is complex and occurs at multiple levels, but phosphorylation of AfIR by protein kinase A inactivates AfIR such that it no longer drives expression of the cluster²⁷. If the activity of a transcription factor that drives expression of an SM cluster is regulated by post-translational modifications, overexpression of the transcription factor may never lead to expression of the genes of the cluster or production of the corresponding SM product.

One theoretical approach to circumventing these problems is to create synthetic transcription factors that are not subject to inhibition by post-translational modification and do not require an unknown natural inducer. To evaluate the efficacy of this approach we have targeted a

silent *A. nidulans* SM biosynthetic gene cluster (Figure 1) containing the highly reducing polyketide synthase gene AN11191 (using the gene designation shared by the Aspergillus Genome Database [AspGD, <http://www.aspergillusgenome.org/>] and Fungal and Oomycete Genomics Resource [Fungidb, <http://fungidb.org/fungidb/>]). Previous efforts to activate this gene cluster by upregulating the transcription factor in the cluster were not successful²². We initially replaced the promoters of two of the genes in the cluster with an inducible promoter, and upregulation of the two genes resulted in the production of (2*Z*,4*Z*,6*E*)-octa-2,4,6-trienoic acid. This compound is of interest as a possible feedstock compound for organic syntheses and as a tanning agent in human skin because it promotes melanogenesis and antioxidant defense in melanocytes²⁸. This result suggested that the final product of the cluster might be of interest.

We now report that we have implemented the hybrid, or synthetic, transcription factor approach successfully to activate this gene cluster. We have added an activation domain from AfoA, the transcription factor that drives expression of the asperfuranone gene cluster, to the transcription factor from the cluster. Upregulation of the hybrid transcription factor results in expression of genes in the silent cluster responsible for the biosynthesis of an anti-inflammatory, antitumor antibiotic, (+)-asperlin^{29–33}. The production of (+)-asperlin was previously reported in *A. nidulans*, *A. caespitosus*, and *A. versicolor*^{29–31,34}, but the specific gene cluster for this polyketide has never been identified.

RESULTS AND DISCUSSION

Induction of AN11199 and AN11191 results in the production of (2*Z*,4*Z*,6*E*)-octa-2,4,6-trienoic acid.

One strategy our labs have used to elicit SM production is to replace the promoters of transcription factor genes associated with SM gene clusters with the inducible *alcA* promoter [*alcA*(p)] followed by induction of expression. We previously found, however, that the replacement of the promoter of AN9221, a transcription factor gene associated with the SM gene cluster containing the highly reducing polyketide synthase gene AN11191, with *alcA*(p) did not result in production of a detectable product²². Another strategy we have used is to simply replace the promoters of SM cluster genes individually with *alcA*(p) and induce expression^{22,23}. We consequently replaced the promoter of AN11191 with the *alcA* promoter using an approach we have previously reported²³. We also replaced the promoter of AN11199, a gene that we hypothesized might be required for release of product from AN11191. The resulting strain was LO4912 (genotypes of strains in Table 1). Induction of expression of these genes resulted in production of compound **1** (2*Z*,4*Z*,6*E*)-octa-2,4,6-trienoic acid (Figure 2) (details of purification and structure elucidation are in the 1.2). (2*Z*,4*Z*,6*E*)-octa-2,4,6-trienoic acid is a particularly interesting compound because of its potential as a natural tanning agent for human skin²⁸. It is a promoter of melanogenesis and antioxidant defense in melanocytes in vitro and in vivo, topically and systemically. Pigmentation is a powerful protectant from UV damage and skin cancer, and octatrienoic acid is an attractive natural photoprotectant. The alternating double bonds in this molecule also suggest that it has potential as a feedstock for organic syntheses. Since only two of the genes of the biosynthetic gene cluster were upregulated, we believed it was likely that this

compound was an intermediate rather than the final product of the biosynthetic pathway, and we wished to develop a method for activating expression of the entire pathway.

Construction and expression of a synthetic hybrid transcription factor drives expression of the AN11191 gene cluster.

Although individual replacement of promoters can be a very powerful tool for eliciting production of SMs²³, there are inherent limitations. The most obvious is that one must carry out several transformations to replace the promoters of all target genes. Our molecular genetic methods are efficient, however, and these transformations can be carried out relatively quickly. Another, more serious, problem is that if any of the genes of the target cluster is incorrectly annotated, promoter replacement for that gene will be non-productive. The gene will not be expressed or only a gene fragment will be expressed. In addition, although progress has been made, defining the boundaries of gene clusters from genomic sequence alone is still an inexact process⁴. We searched, therefore, for a new approach that would allow expression of entire clusters.

One potential approach that might overcome post-translational inactivation of SM transcription factors, and/or Supporting Information sections 1.1 and requirement of an unknown natural inducer for SM transcription factors, is to exploit the domain modularity of transcription factors^{35,36}. Synthetic hybrid transcription activators may be constructed by fusing an active transcription activation domain from a transcription factor known to function in late stationary growth phase to a target SM transcription factor or its DNA binding domain (DBD), which provides the DNA sequence specificity tethering the activation domain to SM gene cluster promoters. There are a number of factors to consider in designing such a synthetic transcription factor. Transcription initiation is complex and different transcription factors may be more or less potent in different contexts. In our experience, SMs are not produced at high levels during log phase growth even if the SM biosynthetic genes are under control of promoters that drive high levels of expression in log phase. The likely explanation is that the precursors required for SM biosynthesis are limiting in log growth. In late stationary phase, however, SM production, particularly polyketide production, can be driven to very high levels, revealing that precursors are plentiful. A suitable transcription factor must be able to drive high levels of transcription at that time. The *afmA* gene drives expression of the asperfuranone biosynthetic cluster²⁵, and upregulation of *afmA* has allowed us to obtain very high levels of production of asperbenzaldehyde, an intermediate in asperfuranone biosynthesis²⁶ (> 2 g/L of purified product). AfoA (the protein encoded by the *afmA* gene), is, thus, capable of driving high levels of expression in late stationary phase cultures when SM precursors are abundant, and it is not negatively regulated post-translationally to any significant extent. We, therefore, chose *afmA* as the source for our activation domain.

We note that in these experiments our replacement of the promoter of *afmA* (AN1029) with the *alcA* promoter, was based on the AspGD version 3 gene annotation (AN1029.3)²⁵. This annotation of AN1029 was identical to the previous version 2 annotation. Upregulation of the *alcA* promoter drove expression of a functional transcription factor that resulted in a successful upregulation of asperfuranone production²⁵ and, with an appropriate pathway

deletion, asperbenzaldehyde production.²⁶ These results indicate that the version 3 annotation was correct, at least with respect to the annotation of the start codon. The current gene annotation for *afmA* in the AspGD and FungiDB databases (<http://www.aspergillusgenome.org>; <http://fungidb.org>) is version 5. In this version *afmA* is shortened at the 5' end relative to the version 3 annotation, omitting exon 1 and the DBD. We do not believe that the version 5 annotation is correct because the DBD would be required for function and because promoter replacements based on the longer, version 3 annotation drive expression of a functional transcription factor. We have consequently based our analyses on the version 3 annotation. We have included the version 3 nucleotide and amino acid sequences for *afmA* and AfoA in the supplemental materials (Supporting Information Figure S1). The identical version 2 annotation (AN1029.2) is still available at the National Center for Biotechnology Information (<https://www.ncbi.nlm.nih.gov/>).

The AfoA transcription factor harbors a Zn(II)2Cys6 DNA binding motif. The DNA-binding specificity of Zn(II)2Cys6 transcription factors is typically provided by the specificity region and heptad repeat coiled-coil forming dimerization motif immediately following the Zn(II)2Cys6 motif, whereas the activation domain is C-terminal to the heptad repeats^{36,40}. The predicted Zn(II)2Cys6 DNA-binding motif is at residues 16-43, and a potential coiled-coil dimerization motif is at residues 55-80. A second predicted coiled coil at 98-129 could contribute to dimerization. A putative nuclear localization signal (NLS) of the SV40 Large T-antigen-type is found at residues 236-242.

To verify that AfoA residues C-terminal to the DBD contained the activation domain and could activate gene expression in a hybrid transcription factor context we replaced amino acids 143-867 of the transcription factor FacB with amino acids 130-666 of AfoA (Figure 3). N-terminal residues 1-142 of FacB are sufficient for DNA binding and are known to provide DNA binding specificity for acetate utilization gene promoters when fused to the activation domain of other Zn(II)2Cys6 transcription factors⁴¹⁻⁴⁴. Fusion of putative ADs to the FacB DBD in strains lacking native *facB* provides a simple assay for activation function assessed by growth on acetate as a sole carbon source. The hybrid FacB(1-142)-AfoA(130-666) transcription factor activated transcription of acetate utilization genes as growth on acetate was supported albeit below wild-type growth rates (Figure 3).

Having established that the *afmA* activation domain can activate expression in the context of a hybrid transcription factor, we attempted to create a hybrid transcription factor that would drive expression of the SM gene cluster containing AN11191 and AN11199. We fused the transcription activation domain (residues 130-666) of AfoA to amino acids 250 through 652 of AN9221, as annotated in the AspGD and FungiDB databases. (We have subsequently found that AN9221 was mis-annotated, as we will discuss.) This region included the AN9221 Zn(II)2Cys6 motif, specificity region, and heptad repeats (residues 269-352) predicted to provide DNA-binding specificity for promoters in the cluster regulated by AN9221 (Figure 4). We omitted AN9221 N-terminal residues 2-249 to avoid predicted transmembrane domains, but retained the AN9221 sequences C-terminal to the DNA-binding domain in the fusion protein as they may contribute to activation, particularly if an intermediate in the regulated pathway acts as an inducer, as is known for some other transcription factors of this class (e.g. LeuB⁴⁵, AmyR⁴⁶). The resultant hybrid transcription

factor, consisting of AN9221 residues 250-652, and AfoA residues 130-666, was put under control of the *alcA* promoter (with a start codon created during the fusion PCR process) and inserted at the *yA* locus using the *Aspergillus terreus pyrG* pyrimidine prototrophic selectable marker (Figure 4). The recipient strain LO9577 is derived from LO8030, which carries deletions of eight of the most highly expressed gene clusters in *A. nidulans* including the asperfuranone biosynthetic gene cluster³⁹. Removal of these gene clusters lowers the SM background and facilitates detection of newly induced metabolites. LO9577 also carries a replacement of the *yA* gene with a fragment containing the *Aspergillus fumigatus pyroA* gene (*AfpyroA*) as well as the *aldA* promoter driving expression of *afoA*. This facilitated construction and identification of correct transformants as shown (Figure 4). In particular, correct transformants with the *alcA*(p)-driven hybrid transcription factor gene targeted in single copy at *yA* and replacing *AfpyroA-aldA*(p)*afoA* will be pyrimidine prototrophs and pyridoxine auxotrophs. Transformants were analyzed by diagnostic PCR, and we subsequently have worked with a correct transformant designated LO9721.

We next wished to determine if the hybrid transcription factor drove expression of the target gene cluster and, if so, whether this led to the production of the SM product of the cluster. We examined expression of cluster genes using RNA-seq. Our samples were grown in liquid culture with lactose minimal medium. Lactose is a non-repressing carbon source that allows expression from the *alcA* promoter. The *alcA* promoter is induced by a number of compounds. For induction of expression of the hybrid transcription factor from the *alcA* promoter in these experiments, we used 50 mM methyl-ethyl ketone (MEK). We carried out RNA-seq (in triplicate) for LO1362, a control strain wild-type for SM gene clusters, LO9721 under non-repressing conditions, which is predicted to give some expression of the hybrid transcription factor, and LO9721 under inducing conditions, which is predicted to give strong expression of the hybrid transcription factor.

We first verified that the hybrid transcription factor was expressed by looking for *afoA* sequence reads. As the wild-type *afoA* gene in the asperfuranone gene cluster has been deleted in LO9721, any *afoA* reads will be from the hybrid transcription factor. *afoA* reads are quantified in Table 2. There are almost no *afoA* reads in the control strain, in which the asperfuranone gene cluster is intact. This reflects the fact that the native *afoA* promoter is essentially completely off in control strains with wild-type regulation of the asperfuranone gene cluster. In LO9721 under non-repressing conditions for the *alcA* promoter, there were moderate numbers of *afoA* reads as expected and induction with MEK substantially boosted the *afoA* reads. We also noted a very large increase in AN9221 reads, consistent with the fact that the hybrid transcription factor includes a portion of AN9221. These data reveal that the hybrid transcription factor is expressed and responds to induction of the *alcA* promoter as expected.

Next, we asked if expression of the hybrid transcription factor resulted in expression of the target gene cluster. Results are shown in Figure 5 and quantified in Table 2. The genes in the region extending from AN11200 through AN11191 are strongly upregulated by expression of the hybrid transcription factor. Adjacent genes are not upregulated. These data reveal that the hybrid transcription factor does, as anticipated, drive expression of the SM cluster genes. They also define the boundaries of the cluster since the genes that are not upregulated are

unlikely to be components of the cluster. We designate this cluster the (+)-asperlin cluster after the product of the cluster (see below) and the genes of this cluster *alnA*-*alnI* plus *alnR* (Figures 1 and 6 and Table 2). Interestingly, differences in the levels of induction of the genes within the cluster is observed. This likely reflects different DNA-binding affinities of the hybrid transcription factor for different *aln* gene promoters due to either sequence differences in the different AlnR DNA-binding sites or different numbers of AlnR DNA-binding sites in the promoters. It is tempting to speculate that these expression differences may represent the optimal relative levels, arising from DNA-binding site evolution, for operation of the pathway and biosynthesis of the final product.

Our analyses of RNA-seq data also reveal that AN11200 is slightly longer than indicated by the AspGD and FungiDB annotations. AN11200 contains an intron and additional sequence upstream of the current annotation. The corrected coding sequence is in the Supporting Information Figure S2. It is also worth noting that induction of the hybrid transcription factor upregulates the antisense strand of AN11200 as well as the sense strand. Finally, although AN11200 and AN11192 are very close, with the termination codon of AN11200 being less than 250 bp from the start codon of AN11192, our RNA-seq data indicate that they are indeed, separate genes

Reannotation of AN9221.

One unexpected outcome of our RNA sequencing was a realization that AN9221 is misannotated in the databases. We noted in a number of RNA-seq experiments that AN9221 is expressed at very low levels in strains that do not carry the hybrid transcription factor. However, it is expressed enough that a clear pattern in the RNA-seq reads emerged. The reads were confined to what was annotated as exon 3. Reads from exons 1 and 2 were almost completely absent. The fact that transcription of exon 3 was apparently regulated independently of exons 1 and 2 suggested that it is a separate gene. We consequently carried out CAGE RNA-seq, which reveals transcription start sites. The results (Supporting Information Figure S3) revealed that there is a distinct transcription start site at the beginning of exon 3. These data, in combination, reveal exon 3 is a separate transcription unit (i.e. gene) from exons 1 and 2. The coding sequences of these genes along with the amino acid sequences of the predicted products are listed in the Supporting Information Figures S4 and S5. The region containing the sequences previously annotated as exons 1 and 2 appears to be a separate gene encoding transmembrane domains that is upregulated by the hybrid transcription factor. These results indicate that the genes formerly called AN9221 need to be given new designations. We designate the exon 3 gene *alnR* (R denoting that as a transcription factor it is a regulatory gene). Its nucleotide and predicted amino acid sequences are shown in Supporting Information Figure S4. The predicted AlnR protein contains a Zn(II)₂Cys₆ zinc binuclear cluster at residues 24-51 and putative heptad repeats at residues 81-108. We designate the former exons 1 and 2 *alnG*. Our RNA-seq data further indicate that the intron in *alnG* is misannotated in the current AspGD and FungiDB databases. The actual intron is shorter, and the correct *alnG* sequence is given in Supporting Information Figure S5.

Reannotation of AN9221 has important consequences for this study. One is that, in all likelihood, the previously published replacement of the promoter of AN9221 with the *alcA* promoter²² was doomed to failure because the transcription factor was not actually being upregulated. Instead, the adjacent transmembrane domain encoding-gene, *alnG*, was likely upregulated. The second is that in our hybrid transcription factor construction, we actually included most of AlnR (residues 5-407), lacking only the four N-terminal amino acids. This raised the possibility that the activation of the cluster was simply due to upregulation of *alnR* and that the AfoA activation domain was not needed. To test this possibility we placed *alnR* under control of the *alcA* promoter and upregulated it. The resulting HPLC profile (Figure 2) showed that, in inducing conditions, overexpression of *alnR* alone was not sufficient to stimulate SM production by the *aln* gene cluster. The hybrid transcription factor, including the AfoA activation domain, therefore, was required for activation of the cluster.

Expression of the hybrid transcription factor drives production of (+)-asperlin (2).

Strong expression of the hybrid transcription factor resulted in the production of compound **2**, (+)-asperlin, which we believe to be the final product of the target gene cluster (Figures 2 and 6). (+)-Asperlin was purified from large-scale cultures using preparative Thin Layer Chromatography, and the structure was confirmed by NMR (details of purification and structure elucidation are in the Supporting Information section 1.1 and 1.2).

Our RNA-seq analysis provided a clear prediction as to the genes involved in (+)-asperlin biosynthesis, but to test the prediction we individually deleted twelve genes (AN11200-AN9213) and examined the effects of the deletions on the biosynthesis of (+)-asperlin (Figure 5). The deletant strains were cultivated under inducing conditions and their metabolite profiles examined by LC/MS. Analysis of the deletants shows that deletions of *alnA*, *alnB*, *alnC*, *alnD*, *alnF*, *alnH* and *alnI* fully eliminated production while deletion of *alnE* and AN9221 (= *alnR* + *alnG*) diminished production. AN9216, AN9215, AN12145, AN9214, AN9213 strains all produced (+)-asperlin in substantial amounts, revealing that they are not functional components of the (+)-asperlin gene cluster. The deletion and RNA-seq data are consistent with each other and, together, reveal that the genes involved in the biosynthesis of (+)-asperlin (**2**) include; *alnA* (AN11191), *alnB* (AN11199), *alnC* (AN11198), *alnD* (AN9218), *alnE* (AN9219), *alnF* (AN9220), *alnG* (AN9221 exons 1-2), *alnR* (AN9221 exon3), *alnH* (AN11192), and *alnI* (AN11200) (Table 2 and Figure 5).

Earlier attempts to upregulate entire silent gene clusters in *A. nidulans* by expressing pathway-specific transcription factors with an inducible promoter have proven to be largely unsuccessful, resulting in the upregulation of less than 20% of gene clusters tested²². To improve upon this approach, we designed a synthetic hybrid transcription factor to drive expression of the AN11191 cluster. By exploiting the domain modularity of transcription factors, we fused the potent transcription AD of AfoA, the asperfuranone cluster transcription factor, with the DBD of AN9221 (AlnR), the transcription factor of the AN11191 cluster, and placed the synthetic construct under the control of the *alcA* promoter. Induction of expression of this hybrid transcription factor activated expression of AN11191-AN11200 and production of the polyketide, (+)-asperlin (**2**). We confirmed that these genes, but not other genes nearby in the genome, are required for (+)-asperlin (**2**) biosynthesis

through a series of gene deletions. Our results also revealed that the putative native transcription factor for this cluster, AN9221 was misannotated previously and is actually two separate genes. However, induction of expression of the reannotated transcription factor, *alnR*, did not result in significant (+)-asperlin (**2**) production. This result demonstrates the value of the hybrid transcription factor approach for activation of expression of the (+)-asperlin (**2**) gene cluster. These results, taken together, show the potential of the synthetic hybrid transcription factor strategy we have developed to upregulate the genes and identify the products of other cryptic clusters.

Cluster-specific transcription factors are often associated with SM gene clusters in Ascomycetes. In *A. nidulans* alone, 48 of 71 putative SM clusters contain identifiable transcription factors³. One-half to two-thirds of putative SM clusters in related *Aspergillus* species contain identifiable transcription factor genes; *Aspergillus oryzae*: 40 of 75, *Aspergillus fumigatus*: 27 of 39, *Aspergillus niger*: 54 of 81, *Aspergillus flavus*: 35 of 55^{3,48}. Moreover, most of these transcription factors belong to the Zn(II)2Cys6 DBD class, and the majority of SM products from these clusters are unknown. This indicates the potential scope of application of hybrid transcription factor technology for SM discovery. The success of our hybrid transcription factor strategy relied on the Zn(II)2Cys6 DBD of *AlnR*. This class of transcription factors usually binds DNA as homodimers^{36,40}. The hybrid transcription factor approach should be applicable to other DBD classes. Hybrid transcription factors that tether an active activation domain to the DBD of a SM cluster transcription factor have the advantage of providing cluster activation specificity without needing to know whether a natural inducer or post-translational modification of the activation domain of the native transcription factor is necessary for activation. In some cases, however, the success of this approach may be hampered by the requirement of an unknown heterodimerization partner or other DNA-binding partner for promoter binding, by post-translational regulation of the DBD of the cluster-specific transcription factor, or by the availability of the necessary metabolic precursors. Future efforts with respect to the (+)-asperlin (**2**) cluster will focus on identifying intermediates of and fully elucidating the stepwise biosynthetic pathway for (+)-asperlin (**2**).

METHODS

Bioinformatic analysis of transcription factors.

The location of predicted structural features of the transcription factors was determined following analysis using ScanProsite (<https://prosite.expasy.org/>)⁴⁹ for the Zn(II)2Cys6 domain, Lupas coiled-coil prediction (<http://toolkit.tuebingen.mpg.de/pcoils>)⁵⁰ for the dimerization motif, and PSORTII (<https://psort.hgc.jp/>)⁵¹ for the Nuclear Localization Signals (NLSs). The transmembrane domains in the mis-annotated AN9221 sequences were analyzed using the TMHMM Server v. 2.0 for prediction of transmembrane helices in proteins (<http://www.cbs.dtu.dk/services/TMHMM/>)⁵².

Fermentation and LC/MS Analysis.

For all strains, 3×10^7 spores were grown in 30 mL lactose minimum media (LMM) (15 g L^{-1} D-lactose, 6 g L^{-1} NaNO_3 , 0.52 g L^{-1} KCl, 0.52 g L^{-1} $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 1.52 g L^{-1}

KH_2PO_4 , and 1 mL L⁻¹ Hutner's trace element solution⁵³) or glucose minimum media (GMM) (10 g L⁻¹ D-glucose, 6 g L⁻¹ NaNO₃, 0.52 g L⁻¹ KCl, 0.52 g L⁻¹ MgSO₄·7H₂O, 1.52 g L⁻¹ KH₂PO₄, and 1 mL L⁻¹ Hutner's trace element solution), and, where appropriate, supplemented with 0.5 mg L⁻¹ pyridoxine, and 2.5 mg L⁻¹ riboflavin, in 125 mL flasks at 37 °C with shaking at 180 rpm. For *alcA(p)* induction, 50 mM of MEK was added to the culture 42 h after inoculation². Culture medium and hyphae were collected 72 h after induction by filtration and extracted twice with a volume of EtOAc equal to the culture volume. The combined EtOAc layers were evaporated *in vacuo*, redissolved in 1 mL of MeOH and 10 μL was injected for HPLC-DAD-MS analysis. LC/MS spectra were obtained using a ThermoFinnigan LCQ Advantage ion trap mass spectrometer with a reverse phase C₁₈ column (Alltech Prevail C₁₈; particle size, 3 μm ; column, 2.1 by 100 mm) at a flow rate of 125 $\mu\text{L min}^{-1}$. The solvent gradient for LC/MS was 95% MeCN–H₂O (solvent B) in 5% MeCN–H₂O (solvent A), both of which contained 0.05% formic acid, as follows: 0% solvent B from 0 to 5 min, 0 to 100% solvent B from 5 min to 35 min, 100% solvent B from 35 to 40 min, 100 to 0% solvent B from 40 to 45 min, and reequilibration with 0% solvent B from 45 to 50 min. Conditions for MS included a capillary voltage of 5.0 kV, a sheath gas flow rate at 60 arbitrary units, an auxiliary gas flow rate at 10 arbitrary units, and the ion transfer capillary temperature at 350 °C.

Compound Spectral Data.

NMR spectral data were collected on a Varian Mercury Plus 400 spectrometer. High resolution electrospray ionization mass spectrum (HRESIMS) was obtained on ThermoScientific Q Exactive hybrid quadrupole-Orbitrap mass spectrometer with an Eclipse XDB-C₁₈ column (Agilent 5 μm 4.6 \times 150 mm) at a flow rate of 125 $\mu\text{L min}^{-1}$. Conditions for MS included a spray voltage of 3.5 kV, sheath gas flow rate 20 au, auxiliary gas flow rate at 5 au, sweep gas flow rate at 1 au, capillary temperature at 275 °C, s-lens RF level 55, auxiliary gas heat temperature at 325 °C, scan range of 100-600 m/z, resolution 140,000, AGC target 3 \times 10⁶, and maximum injection time of 200 ms.

(2Z,4Z,6E)-octa-2,4,6-trienoic acid (1): White amorphous powder; for UV-Vis and HRESIMS spectra see Supporting Information Figure S6; for ¹H and ¹³C NMR data (CDCl₃), see Supporting Information Figure S7-S9; HRESIMS, [M – H]⁺ *m/z* found 139.0753.

(+)-Asperlin (2): Colorless oil; for UV-Vis and HRESIMS spectra see Supporting Information Figure S6; for ¹H and ¹³C NMR data (CDCl₃), see Supporting Information Figure S7, S10-S11; HRESIMS, [M – H]⁺ *m/z* found 213.0754.

Transformation procedures and construction of transforming molecules.

Transforming molecules were constructed by fusion PCR as previously described^{37,54,55}. The selectable markers used were the *Aspergillus fumigatus pyrG* gene (*AfpyrG*)⁵⁶, the *Aspergillus terreus pyrG* gene (*AtpyrG*)⁵⁷, the *A. fumigatus riboB* gene *AriboB*⁵¹ and the *A. fumigatus pyroA* gene⁵¹. The markers used for each construct are specified in the strain list. All constructs were verified by diagnostic PCR. Transformation procedures were as previously described^{54,55}.

RNA-seq.

Conidia of a control strain (LO1362 = TNO2A7) and of LO9721, which carries the hybrid transcription factor under control of the *alcA* promoter, were inoculated into 50 mL of lactose minimal medium in 250 mL Erlenmeyer flasks at a density of 2×10^4 conidia mL⁻¹. The medium was supplemented with riboflavin, pyridoxine, uridine and uracil to supplement the nutritional markers carried by the control strain. The flasks were incubated at 37 °C in a rotary shaker at 100 RPM. After 36 h incubation, MEK was added to a final concentration of 50 mM to flasks containing the control strain and to half the flasks containing LO9721. The remaining LO9721 flasks were incubated without MEK. After four hours of further incubation, hyphae were harvested by filtration through miracloth. They were then frozen with liquid nitrogen and reduced to a frozen powder with a Multi-beads Shocker (Yasui Kikai Co., Ltd., Japan). Library preparation, sequencing and data analysis procedures were as previously described²⁰ with minor modifications (for example, HISAT2 was used instead of HISAT for read mapping).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENT

Supported by the National Institutes of Health (grant number P01GM084077), by US Department of Energy funds awarded through Pacific Northwest National Laboratory, by the Irving S. Johnson Fund of the Kansas University Endowment Association, and by the Kansas City Area Life Sciences Institute and Hall Family Foundation (grant number KCALSI 17-12). Contribution no. 19-004-J from the Kansas Agricultural Experiment Station. The authors would like to thank O. Yoon and C.B. Jenkinson (University of Kansas) for technical assistance. We thank N.A. Graham and A. Delfarah (USC Viterbi School of Engineering) for their help collecting HRESIMS data.

REFERENCES

- (1). Keller NP, Turner G, and Bennett JW (2005) Fungal secondary metabolism—from biochemistry to genomics. *Nat. Rev. Microbiol* 3, 937–947. [PubMed: 16322742]
- (2). Galagan JE, Calvo SE, Cuomo C, Ma L-J, Wortman JR, Batzoglou S, Lee S-I, Ba türkmen M, Spevak CC, Clutterbuck J, Kapitonov V, Jurka J, Scacciocchio C, Farman M, Butler J, Purcell S, Harris S, Braus GH, Draht O, Busch S, D’Enfert C, Bouchier C, Goldman GH, Bell-Pedersen D, Griffiths-Jones S, Doonan JH, Yu J, Vienken K, Pain A, Freitag M, Selker EU, Archer DB, Peñalva MÁ, Oakley BR, Momany M, Tanaka T, Kumagai T, Asai K, Machida M, Nierman WC, Denning DW, Caddick M, Hynes M, Paoletti M, Fischer R, Miller B, Dyer P, Sachs MS, Osmani SA, and Birren BW (2005) Sequencing of *Aspergillus nidulans* and comparative analysis with *A. fumigatus* and *A. oryzae*. *Nature* 438, 1105–1115. [PubMed: 16372000]
- (3). Inglis DO, Binkley J, Skrzypek MS, Arnaud MB, Cerqueira GC, Shah P, Wymore F, Wortman JR, and Sherlock G (2013) Comprehensive annotation of secondary metabolite biosynthetic genes and gene clusters of *Aspergillus nidulans*, *A. fumigatus*, *A. niger* and *A. oryzae*. *BMC Microbiol.* 13, 91. [PubMed: 23617571]
- (4). de Vries RP, Riley R, Wiebenga A, Aguilar-Osorio G, Amillis S, Uchima CA, Anderluh G, Asadollahi M, Askin M, Barry K, Battaglia E, Bayram Ö, Benocci T, Braus-Stromeyer SA, Caldana C, Cánovas D, Cerqueira GC, Chen F, Chen W, Choi C, Clum A, dos Santos RAC, Damásio AR, de L, Diallinas G, Emri T, Fekete E, Flippi M, Freyberg S, Gallo A, Gournas C, Habgood R, Hainaut M, Harispe ML, Henrissat B, Hildén KS, Hope R, Hossain A, Karabika E, Karaffa L, Karányi Z, Kraševc N, Kuo A, Kusch H, LaButti K, Legendijk EL, Lapidus A, Lévassieur A, Lindquist E, Lipzen A, Logrieco AF, MacCabe A, Mäkelä MR, Malavazi I, Melin P, Meyer V, Mielnichuk N, Miskei M, Molnár ÁP, Mulé G, Ngan CY, Orejas M, Orosz E,

Ouedraogo JP, Overkamp KM, Park H-S, Perrone G, Piumi F, Punt PJ, Ram AFJ, Ramón A, Rauscher S, Record E, Riaño-Pachón DM, Robert V, Röhrig J, Ruller R, Salamov A, Salih NS, Samson RA, Sándor E, Sanguinetti M, Schütze T, Sepi K, Shelest E, Sherlock G, Sophianopoulou V, Squina FM, Sun H, Susca A, Todd RB, Tsang A, Unkles SE, van de Wiele N, van Rossen-Uffink D, Oliveira JV de C, Vesth TC, Visser J, Yu J-H, Zhou M, Andersen MR, Archer DB, Baker SE, Benoit I, Brakhage AA, Braus GH, Fischer R, Frisvad JC, Goldman GH, Houbraeken J, Oakley B, Pócsi I, Scazzocchio C, Seiboth B, vanKuyk PA, Wortman J, Dyer PS, and Grigoriev IV (2017) Comparative genomics reveals high biological diversity and specific adaptations in the industrially and medically important fungal genus *Aspergillus*. *Genome Biol.* 18.

- (5). Nielsen JC, Grijseels S, Prigent S, Ji B, Dainat J, Nielsen KF, Frisvad JC, Workman M, and Nielsen J (2017) Global analysis of biosynthetic gene clusters reveals vast potential of secondary metabolite production in *Penicillium* species. *Nat. Microbiol* 2, 17044. [PubMed: 28368369]
- (6). Brakhage AA, Schuemann J, Bergmann S, Scherlach K, Schroeckh V, and Hertweck C (2008) Activation of fungal silent gene clusters: A new avenue to drug discovery, in *Natural Compounds as Drugs*, pp 1–12. Birkhäuser Basel.
- (7). Brakhage AA, and Schroeckh V (2011) Fungal secondary metabolites – Strategies to activate silent gene clusters. *Fungal Genet. Biol.* 48, 15–22. [PubMed: 20433937]
- (8). Chiang Y-M, Chang S-L, Oakley BR, and Wang CC (2011) Recent advances in awakening silent biosynthetic gene clusters and linking orphan clusters to natural products in microorganisms. *Curr. Opin. Chem. Biol* 15, 137–143. [PubMed: 21111669]
- (9). Lim FY, Sanchez JF, Wang CCC, and Keller NP (2012) Toward awakening cryptic secondary metabolite gene clusters in filamentous fungi, in *Methods in Enzymology*, pp 303–324. Elsevier.
- (10). Brakhage AA (2013) Regulation of fungal secondary metabolism. *Nat. Rev. Microbiol* 11, 21–32. [PubMed: 23178386]
- (11). Yaegashi J, Oakley BR, and Wang CCC (2014) Recent advances in genome mining of secondary metabolite biosynthetic gene clusters and the development of heterologous expression systems in *Aspergillus nidulans*. *J. Ind. Microbiol. Biotechnol.* 41, 433–442. [PubMed: 24342965]
- (12). Andersen MR, Nielsen JB, Klitgaard A, Petersen LM, Zachariassen M, Hansen TJ, Blicher LH, Gotfredsen CH, Larsen TO, Nielsen KF, and Mortensen UH (2013) Accurate prediction of secondary metabolite gene clusters in filamentous fungi. *Proc. Natl. Acad. Sci* 110, E99–E107. [PubMed: 23248299]
- (13). Bok JW, Chiang Y-M, Szewczyk E, Reyes-Dominguez Y, Davidson AD, Sanchez JF, Lo H-C, Watanabe K, Strauss J, Oakley BR, Wang CCC, and Keller NP (2009) Chromatin-level regulation of biosynthetic gene clusters. *Nat. Chem. Biol* 5, 462–464. [PubMed: 19448638]
- (14). Soukup AA, Chiang Y-M, Bok JW, Reyes-Dominguez Y, Oakley BR, Wang CCC, Strauss J, and Keller NP (2012) Overexpression of the *Aspergillus nidulans* histone 4 acetyltransferase EsaA increases activation of secondary metabolite production. *Mol. Microbiol* 86, 314–330. [PubMed: 22882998]
- (15). Shwab EK, Bok JW, Tribus M, Galehr J, Graessle S, and Keller NP (2007) Histone deacetylase activity regulates chemical diversity in *Aspergillus*. *Eukaryot. Cell* 6, 1656–1664. [PubMed: 17616629]
- (16). Cichewicz RH (2010) Epigenome manipulation as a pathway to new natural product scaffolds and their congeners. *Nat Prod Rep* 27, 11–22. [PubMed: 20024091]
- (17). Vervoort HC, Drašković M, and Crews P (2011) Histone deacetylase inhibitors as a tool to up-regulate new fungal biosynthetic products: isolation of EGM-556, a cyclodepsipeptide, from *Microascus* sp. *Org. Lett* 13, 410–413. [PubMed: 21174394]
- (18). Albright JC, Henke MT, Soukup AA, McClure RA, Thomson RJ, Keller NP, and Kelleher NL (2015) Large-scale metabolomics reveals a complex response of *Aspergillus nidulans* to epigenetic perturbation. *ACS Chem. Biol* 10, 1535–1541. [PubMed: 25815712]
- (19). Keller N, Bok J, Chung D, Perrin RM, and Keats Shwab E (2006) LaeA, a global regulator of *Aspergillus* toxins. *Med. Mycol* 44, 83–85.
- (20). Oakley CE, Ahuja M, Sun W-W, Entwistle R, Akashi T, Yaegashi J, Guo C-J, Cerqueira GC, Wortman JR, Wang CCC, Chiang Y-M, and Oakley BR (2017) Discovery of McrA, a master

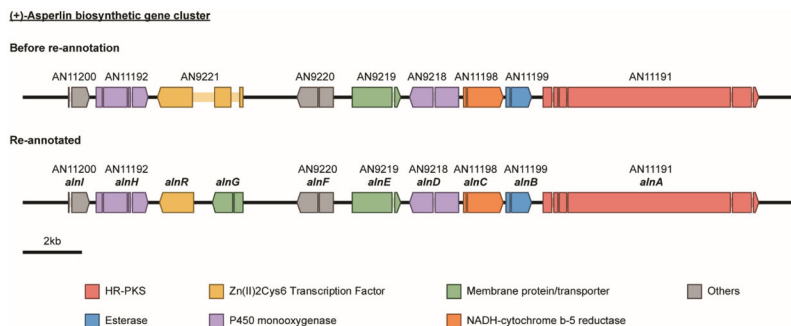
regulator of *Aspergillus* secondary metabolism. *Mol. Microbiol* 103, 347–365. [PubMed: 27775185]

- (21). Schroeckh V, Scherlach K, Nutzmann H-W, Shelest E, Schmidt-Heck W, Schuemann J, Martin K, Hertweck C, and Brakhage AA (2009) Intimate bacterial-fungal interaction triggers biosynthesis of archetypal polyketides in *Aspergillus nidulans*. *Proc. Natl. Acad. Sci* 106, 14558–14563. [PubMed: 19666480]
- (22). Ahuja M, Chiang Y-M, Chang S-L, Praseuth MB, Entwistle R, Sanchez JF, Lo H-C, Yeh H-H, Oakley BR, and Wang CCC (2012) Illuminating the diversity of aromatic polyketide synthases in *Aspergillus nidulans*. *J. Am. Chem. Soc* 134, 8212–8221. [PubMed: 22510154]
- (23). Yeh H-H, Ahuja M, Chiang Y-M, Oakley CE, Moore S, Yoon O, Hajovsky H, Bok J-W, Keller NP, Wang CCC, and Oakley BR (2016) Resistance gene-guided genome mining: serial promoter exchanges in *Aspergillus nidulans* reveal the biosynthetic pathway for fellutamide B, a proteasome inhibitor. *ACS Chem. Biol* 11, 2275–2284. [PubMed: 27294372]
- (24). Bergmann S, Schumann J, Scherlach K, Lange C, Brakhage AA, and Hertweck C (2007) Genomics-driven discovery of PKS-NRPS hybrid metabolites from *Aspergillus nidulans*. *Nat. Chem. Biol* 3, 213–217. [PubMed: 17369821]
- (25). Chiang Y-M, Szewczyk E, Davidson AD, Keller N, Oakley BR, and Wang CCC (2009) A gene cluster containing two fungal polyketide synthases encodes the biosynthetic pathway for a polyketide, asperfuranone, in *Aspergillus nidulans*. *J. Am. Chem. Soc* 131, 2965–2970. [PubMed: 19199437]
- (26). Paranjape SR, Riley AP, Somoza AD, Oakley CE, Wang CCC, Prisinzano TE, Oakley BR, and Gamblin TC (2015) Azaphilones inhibit Tau aggregation and dissolve Tau aggregates *in Vitro*. *ACS Chem. Neurosci* 6, 751–760. [PubMed: 25822288]
- (27). Shimizu K, Hicks JK, Huang T-P, and Keller NP (2003) Pka, Ras and RGS protein interactions regulate activity of AflR, a Zn(II)2Cys6 transcription factor in *Aspergillus nidulans*. *Genetics* 165, 1095–1104. [PubMed: 14668367]
- (28). Flori E, Mastrofrancesco A, Kovacs D, Ramot Y, Briganti S, Bellei B, Paus R, and Picardo M (2011) 2,4,6-Octatrienoic acid is a novel promoter of melanogenesis and antioxidant defense in normal human melanocytes via PPAR- γ activation. *Pigment Cell Melanoma Res.* 24, 618–630. [PubMed: 21762468]
- (29). Argoudelis AD, Coats JH, and Herr RR (1965) Isolation and characterization of a new antibiotic, U-13,933. *Antimicrob. Agents Chemother.* 5, 801–803. [PubMed: 5883500]
- (30). Argoudelis AD, and Ziesler JF (1966) The structure of U-13,933, a new antibiotic. *Tetrahedron Lett.* 7, 1969–1973.
- (31). Mizuba S, Lee K, and Jiu J (1975) Three antimicrobial metabolites from *Aspergillus caespitosus*. *Can. J. Microbiol.* 21, 1781–1787. [PubMed: 1104120]
- (32). Lee D-S, Jeong G-S, Li B, Lee SU, Oh H, and Kim Y-C (2011) Asperlin from the marine-derived fungus *Aspergillus* sp. SF-5044 exerts anti-inflammatory effects through Heme Oxygenase-1 expression in murine macrophages. *J. Pharmacol. Sci* 116, 283–295. [PubMed: 21705844]
- (33). He L, Nan M-H, Oh HC, Kim YH, Jang JH, Erikson RL, Ahn JS, and Kim BY (2011) Asperlin induces G2/M arrest through ROS generation and ATM pathway in human cervical carcinoma cells. *Biochem. Biophys. Res. Commun* 409, 489–493. [PubMed: 21600879]
- (34). Zhou Y, Chen R, Liu D, Wu C, Guo P, and Lin W (2017) Asperlin inhibits LPS-evoked foam cell formation and prevents atherosclerosis in ApoE $^{-/-}$ mice. *Mar. Drugs* 15, 358.
- (35). Keegan L, Gill G, and Ptashne M (1986) Separation of DNA binding from the transcription-activating function of a eukaryotic regulatory protein. *Science* 231, 699–704. [PubMed: 3080805]
- (36). Reece RJ, and Ptashne M (1993) Determinants of binding-site specificity among yeast C6 zinc cluster proteins. *Science* 261, 909–911. [PubMed: 8346441]
- (37). Nayak T, Szewczyk E, Oakley CE, Osmani A, Ukil L, Murray SL, Hynes MJ, Osmani SA, and Oakley BR (2006) A versatile and efficient gene-targeting system for *Aspergillus nidulans*. *Genetics* 172, 1557–1566. [PubMed: 16387870]
- (38). Chiang Y-M, Szewczyk E, Davidson AD, Entwistle R, Keller NP, Wang CCC, and Oakley BR (2010) Characterization of the *Aspergillus nidulans* monodictyphenone gene cluster. *Appl. Environ. Microbiol.* 76, 2067–2074. [PubMed: 20139316]

- (39). Chiang Y-M, Ahuja M, Oakley CE, Entwistle R, Zutz C, Wang CCC, and Oakley BR (2016) Development of genetic dereplication strains in *Aspergillus nidulans* results in the discovery of aspercryptin. *Angew. Chem. Int. Ed Engl* 55, 1662–1665. [PubMed: 26563584]
- (40). Todd RB, and Andrianopoulos A (1997) Evolution of a fungal regulatory gene family: the Zn(II)₂Cys₆ binuclear cluster DNA binding motif. *Fungal Genet. Biol* 21, 388–405. [PubMed: 9290251]
- (41). Todd RB, Murphy RL, Martin HM, Sharp JA, Davis MA, Katz ME, and Hynes MJ (1997) The acetate regulatory gene *facB* of *Aspergillus nidulans* encodes a Zn(II)₂Cys₆ transcriptional activator. *Mol. Gen. Genet* 254, 495–504. [PubMed: 9197408]
- (42). Parsons LM, Davis MA, and Hynes MJ (1992) Identification of functional regions of the positively acting regulatory gene *amdR* from *Aspergillus nidulans*. *Mol. Microbiol* 6, 2999–3007. [PubMed: 1479891]
- (43). Todd RB, Andrianopoulos A, Davis MA, and Hynes MJ (1998) FacB, the *Aspergillus nidulans* activator of acetate utilization genes, binds dissimilar DNA sequences. *EMBO J.* 17, 2042–2054. [PubMed: 9524126]
- (44). Small AJ, Hynes MJ, and Davis MA (1999) The TamA protein fused to a DNA-binding domain can recruit AreA, the major nitrogen regulatory protein, to activate gene expression in *Aspergillus nidulans*. *Genetics* 153, 95–105. [PubMed: 10471703]
- (45). Downes DJ, Davis MA, Kreutzberger SD, Taig BL, and Todd RB (2013) Regulation of the NADP-glutamate dehydrogenase gene *gdhA* in *Aspergillus nidulans* by the Zn(II)₂Cys₆ transcription factor LeuB. *Microbiology* 159, 2467–2480. [PubMed: 24025604]
- (46). Makita T, Katsuyama Y, Tani S, Suzuki H, Kato N, Todd RB, Hynes MJ, Tsukagoshi N, Kato M, and Kobayashi T (2009) Inducer-dependent nuclear localization of a Zn(II)₂Cys₆ transcriptional activator, AmyR, in *Aspergillus nidulans*. *Biosci. Biotechnol. Biochem* 73, 391–399. [PubMed: 19202286]
- (47). Todd RB, Kelly JM, Davis MA, and Hynes MJ (1997) Molecular characterization of mutants of the acetate regulatory gene *facB* of *Aspergillus nidulans*. *Fungal Genet. Biol* 22, 92–102. [PubMed: 9367656]
- (48). Chang P-K, and Ehrlich KC (2013) Genome-wide analysis of the Zn(II)₂Cys₆ zinc cluster-encoding gene family in *Aspergillus flavus*. *Appl. Microbiol. Biotechnol* 97, 4289–4300. [PubMed: 23563886]
- (49). de Castro E, Sigrist CJA, Gattiker A, Bulliard V, Langendijk-Genevaux PS, Gasteiger E, Bairoch A, and Hulo N (2006) ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. *Nucleic Acids Res.* 34, W362–W365. [PubMed: 16845026]
- (50). Biegert A, Mayer C, Remmert M, Soding J, and Lupas AN (2006) The MPI Bioinformatics Toolkit for protein sequence analysis. *Nucleic Acids Res.* 34, W335–W339. [PubMed: 16845021]
- (51). Nakai K, and Horton P (1999) PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization. *Trends Biochem. Sci.* 24, 34–35. [PubMed: 10087920]
- (52). Krogh A, Larsson B, von Heijne G, and Sonnhammer EL. (2001) Predicting transmembrane protein topology with a hidden markov model: application to complete genomes *J. Mol. Biol* 305, 567–580. [PubMed: 11152613]
- (53). Vishniac W, and Santer M (1957) The Thiobacilli. *Bacteriol. Rev* 21, 195–213. [PubMed: 13471458]
- (54). Szweczyk E, Nayak T, Oakley CE, Edgerton H, Xiong Y, Taheri-Talesh N, Osmani SA, and Oakley BR (2007) Fusion PCR and gene targeting in *Aspergillus nidulans*. *Nat. Protoc* 1, 3111–3120.
- (55). Oakley CE, Edgerton-Morgan H, and Oakley BR (2012) Tools for manipulation of secondary metabolism pathways: rapid promoter replacements and gene deletions in *Aspergillus nidulans*, in *Fungal Secondary Metabolism* (Keller NP, and Turner G, Eds.), pp 143–161. Humana Press, Totowa, NJ.
- (56). Weidner G, d'Enfert C, Koch A, Mol PC, and Brakhage AA (1998) Development of a homologous transformation system for the human pathogenic fungus *Aspergillus fumigatus*

based on the *pyrG* gene encoding orotidine 5'-monophosphate decarboxylase. *Curr. Genet.* 33, 378–385. [PubMed: 9618589]

- (57). Dohn JW, Grubbs AW, Oakley CE, and Oakley BR (2018) New multi-marker strains and complementing genes for *Aspergillus nidulans* molecular biology. *Fungal Genet. Biol* 111, 1–6. [PubMed: 29309843]

**Figure 1.**

Organization of the (+)-asperlin biosynthetic gene cluster in *A. nidulans*. The original annotation of the genes in this pathway is included as the top cluster. RNA-seq and CAGE RNA-seq data revealed that AN9221 was incorrectly annotated. We designate the AN9221-exons 1-2 gene as *alnG* (see Supporting Information Figure S5 for the correctly annotated *alnG* sequence), and designate the AN9221-exon 3 gene, encoding a Zn(II)2Cys6 binuclear cluster, as the transcription factor gene *alnR* (see Supporting Information Figure S4 for the correctly annotated *alnG* sequence). The cluster below represents the re-annotated (+)-asperlin biosynthetic gene cluster with the *aln* gene designations included. Predicted gene functions are color-coded.

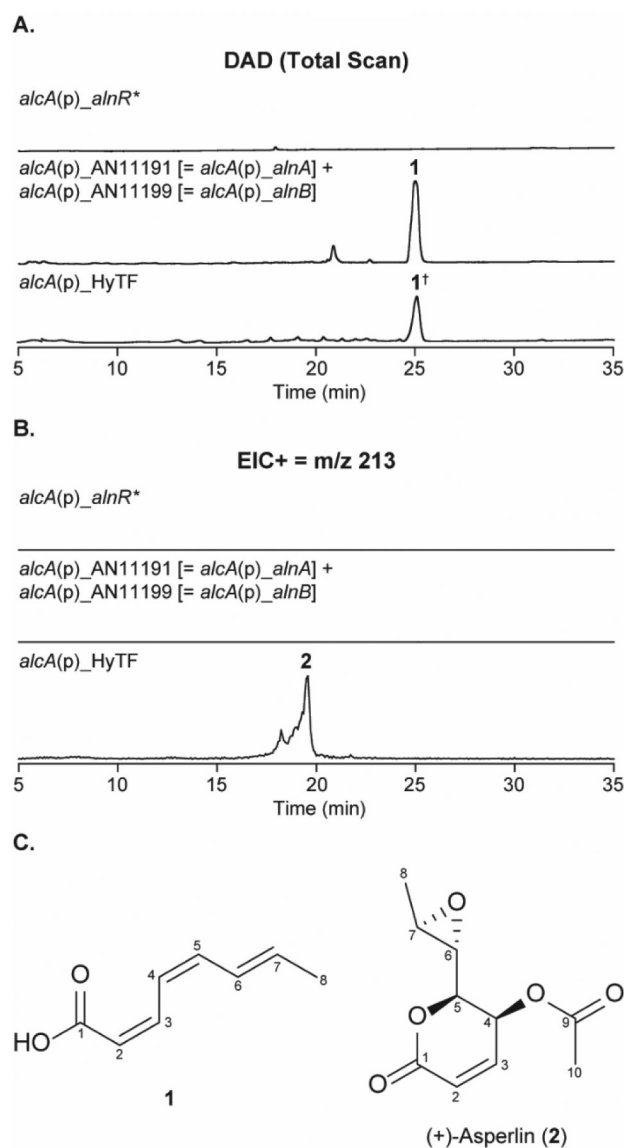


Figure 2. (A) Diode array detector (DAD) total scan HPLC profiles and (B) extracted ion chromatograms (EIC) for compound **2** (EIC = m/z 213) of culture media extracts of *A. nidulans* strains in which we induced the *alcA* promoter driving expression of the native cluster-specific transcription factor gene (*alnR*) (top), the HR-PKS (AN11191 = *alnA*) and esterase (AN11199 = *alnB*) genes (middle), and the hybrid *alnR-afmA* transcription factor gene (HyTF) (bottom) under the control of the *alcA* promoter. **1** is (2*Z*,4*Z*,6*E*)-octa-2,4,6-trienoic acid, and **2** is (+)-asperlin. (C) Chemical structures of **1** and **2** as determined by NMR. *From RNA-seq experiments, we have corrected the originally annotated sequence of the AN9221 transcription factor gene and put the reannotated sequence (*alnR*) under the control of the *alcA* promoter. †At the same retention time as **1**, trace amounts of unknown metabolites could also be detected in the hybrid transcription factor-expressing strain.

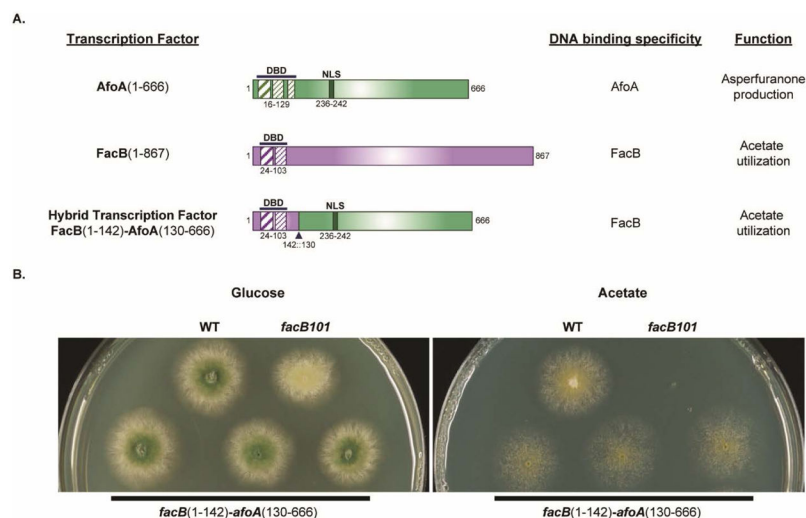


Figure 3. Activation capacity of AfoA residues 130-666 in a hybrid transcription factor context. (A) The primary structure of the transcription factors AfoA and FacB, and the hybrid FacB(1-142)-AfoA(130-666) transcription factor are shown with the relevant functional domains. The DNA Binding Domain (DBD) consists of the Zn(II)₂Cys₆ zinc binuclear cluster (wide diagonal stripe) and the heptad repeats coiled-coil dimerization region (narrow diagonal stripe). The putative nuclear localization signal (NLS) in AfoA is indicated. The activation domain, which is located in the C-terminal portion for this class of transcription factor, is not precisely defined but is located within AfoA residues 130-666. The FacB activation domain is C-terminal to residue 625⁴¹. AfoA sequences are shown in green, whereas FacB sequences are shown in purple. The DNA binding specificity and function of each transcription factor is indicated. The fusion point in the hybrid transcription factor is indicated by an arrow head with the amino acid coordinates separated by two colons. (B) The ability of the FacB(1-142)-AfoA(130-666) hybrid transcription factor to activate FacB-dependent acetate utilization genes was assessed qualitatively by gene replacement of wild type *facB* with the hybrid transcription factor gene containing sequences encoding the FacB DNA binding domain (codons 1-142) and the AfoA activation domain (codons 130 – 666), followed by growth on acetate as a sole carbon source. Growth (37°C, 2 d) of transformants expressing FacB(1-142)-AfoA(130-666) (LO9886, LO9887, and LO9888) are shown relative to the wild type (LO4389) and the *facB101* complete loss-of-function mutant (FGSC 442), which truncates FacB at residue 619 deleting the activation domain⁴⁷. Growth on glucose is independent of *facB* activity. Note that the color difference observed for FGSC 442 is due to the *chaA1* conidial color mutation in the background genotype and is unrelated to *facB*.

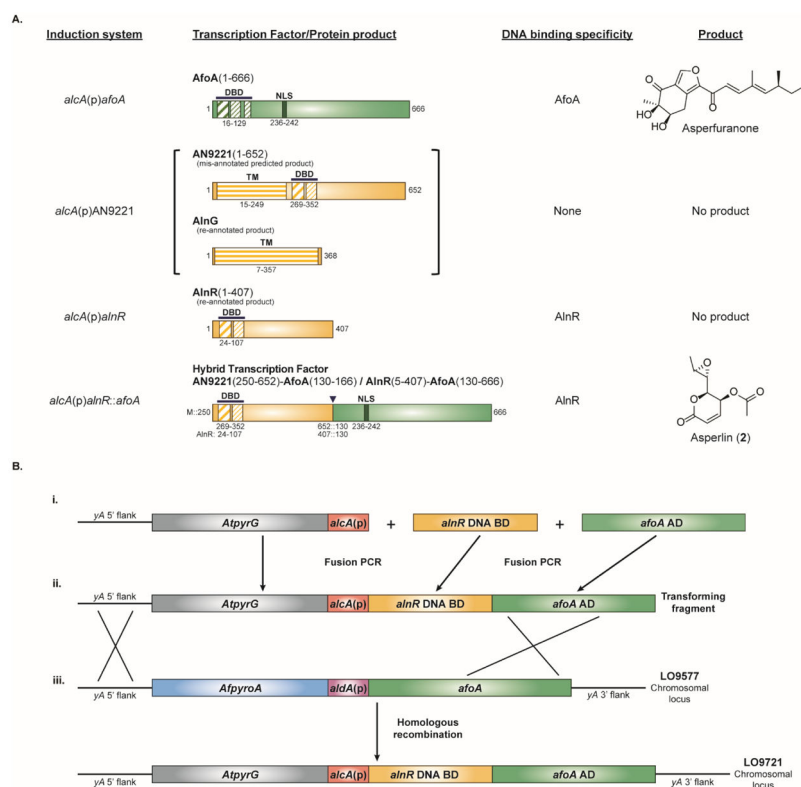


Figure 4. The hybrid AlnR(5-407)-AfoA(130-666) transcription factor leads to (+)-asperlin production.

(A) The induction system used, transcription factor domain structure, DNA binding specificity, and product observed in inducing conditions are summarized from previous experiments and this work. Expression of the AfoA transcription factor from the *alcA(p)* promoter conferred asperfuranone production²⁵, whereas expression of AN9221 from the *alcA(p)* promoter failed to yield a SM product²². The protein expressed from the AN9221 sequence was, due to misannotation, not the predicted transcription factor but instead was the transmembrane (TM) protein AlnG, which lacks a DBD. Expression of the reannotated transcription factor AlnR, which contains the AN9221 DBD, from *alcA(p)* also failed to produce a SM product. Expression of the hybrid AN9221(250-652)-AfoA(130-166) transcription factor, which was reannotated as AlnR(5-407)-AfoA(130-666), conferred production of (+)-asperlin (2). The transcription factors/proteins products are shown with the relevant functional domains: the DBD consists of the Zn(II)₂Cys₆ zinc binuclear cluster (wide diagonal stripe) and the heptad repeats coiled-coil dimerization region (narrow diagonal stripe), the putative NLS in AfoA, and transmembrane region (TM). The activation domain is not precisely defined but is located within AfoA residues 130-666. No obvious NLS is found in AlnR. AfoA sequences are shown in green, whereas AN9221-derived sequences including AlnG and AlnR are shown in yellow. The fusion point in the hybrid transcription factor is indicated by an arrow head with the amino acid coordinates separated by two colons. (B) i. A transforming fragment was created by fusing three DNA fragments using fusion PCR. The first fragment consists of 5' flanking DNA from the *yA* locus, the *Aspergillus terreus pyrG* gene (*AtpyrG*) and the promoter from the alcohol dehydrogenase gene [*alcA(p)*]. This fragment was a cassette previously made by fusion PCR. The second

fragment consists of the DNA binding domain from AN9221 (AN9221 DBD) and the third fragment consists of the activation domain of *afoA* (AN1029) (*afoA* AD). ii. The three fragments were fused by fusion PCR creating the transforming fragment shown. The recipient strain (LO9577) carried a replacement of the *yA* locus by the *Aspergillus fumigatus pyroA* gene (*AfpyroA*), the aldehyde dehydrogenase promoter [*aldA*(p)] and the asperfuranone gene cluster transcription factor, *afoA* (AN1029). iii. Homologous recombination during transformation results in a hybrid transcription factor driven by the *alcA* promoter. Note that *AfpyroA* is replaced by *AtpyrG* resulting in a strain that does not require added pyrimidines for growth but does require added pyridoxine.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

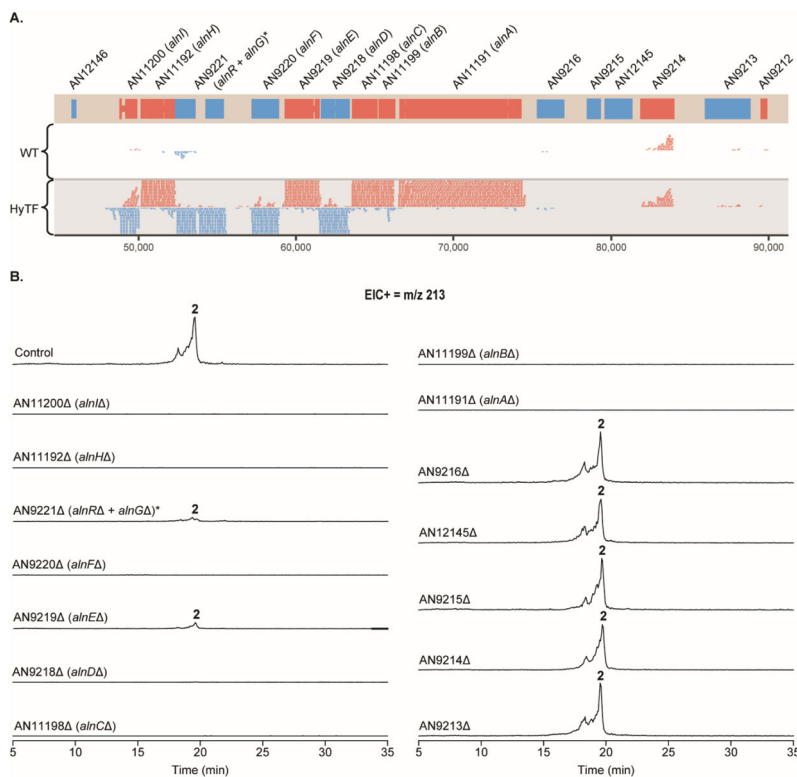
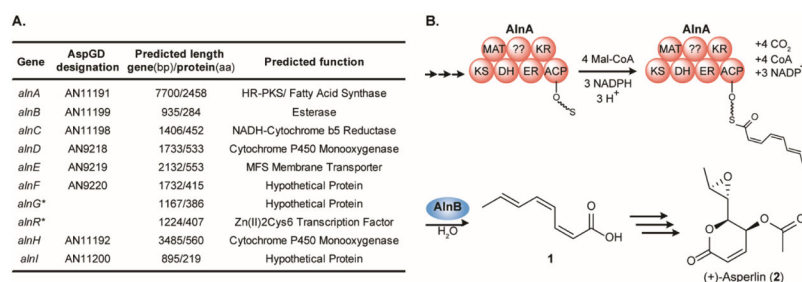


Figure 5.

(A) RNA-seq transcript read alignments to genes surrounding the HR-PKS, AN11191, for the hybrid AlnR-AfoA transcription factor (HyTF) expressing *A. nidulans* strain compared to a strain lacking a hybrid transcription factor. Gene color indicates direction of transcription; genes transcribed from right to left (blue), and genes transcribed from left to right (orange). (B) Extracted ion chromatograms for compound **2** (EIC = *m/z* 213) of culture media extracted from *A. nidulans* hybrid transcription factor expressing strains carrying a single gene deletion of genes surrounding AN11191. *Original annotation of AN9221 suggested 3 exons, however, RNA-seq data indicates that AN9221 is actually two separate transcription units; one gene contains exons 1 and 2 and the other gene contains exon 3.

**Figure 6.**

(A) The deduced functions of each ORF based on the AspGD gene designations. (B) Proposed biosynthesis of *(2Z,4Z,6E)*-octa-2,4,6-trienoic acid (**1**) by AlnA with cleavage of the PKS product by AlnB, and downstream production of the pathway product, (+)-asperlin (**2**). *Original annotation of AN9221 suggested 3 exons, however, RNA-seq data indicates that AN9221 is actually two separate transcription units; *alnG* contains exon 1-2 and the other gene, *alnR*, contains exon 3. KS, β -ketoacyl synthase; MAT, malonyl-CoA:acyl carrier protein transacylase; DH, dehydratase; ER, enoylreductase; KR, ketoreductase; ACP, acyl carrier protein

Table 1.

Strains used in this study

Strains	Genotype	Reference
FGSC442	<i>facB101, riboB2, chaA1, sE15, ntrA14</i>	Fungal Genetics Stock Center
TN02A7 (LO1362)	<i>pyroA4, riboB2, pyrG89, nukA::argB</i>	[37]
LO2026	<i>stcJ ::AfriboB</i> in LO1362	[13], [38]
LO3632	AN9221(<i>alnG-alnR</i>):: <i>AfpyrG-alcA</i> (p)-AN9221 in LO2026	[22]
LO4054	<i>alnA::AfpyrG-alcA</i> (p) <i>alnA</i> in LO2026	This work
LO4387	<i>stcW ::AfpyrG</i> in LO2026	This work
LO4389	<i>stc</i> (AN7804-AN7825) in LO4387	[22]
LO4912	<i>alnB::AfpyroA-alcA</i> (p) <i>alnB</i> in LO4054	This work
LO8030	<i>pyroA4, riboB2, pyrG89, nukA::argB, stc</i> (AN7804-AN7825) , <i>eas</i> (AN2545-AN2549) , <i>atf</i> (ANI1039-ANI1029) , <i>mdp</i> (ANI10023-ANI10021) , <i>tdf</i> (AN8512-AN8520) , <i>aus</i> (AN8379-AN8384, AN9246-AN9259) , <i>oss</i> (AN7906-AN7915) , <i>apt</i> (AN6000-AN6002)	[39]
LO9577	<i>yA::AfpyroA-aldA</i> (p) <i>af</i> <i>oA</i> in LO8030	This work
LO9721	<i>yA::AtpyrG-alcA</i> (p) <i>alnR</i> (250-652)- <i>af</i> <i>oA</i> (130-666) * in LO8030	This work
LO9886, LO9887, LO9888	<i>facB::facB</i> (1-142)- <i>af</i> <i>oA</i> (130-666)- <i>AtpyrG</i> in LO4389	This work
LO11147	<i>yA::AtpyrG-alcA</i> (p)AN9221 in LO9577	This work
LO11280	<i>yA::AfpyrG-alcA</i> (p) <i>alnR</i> in LO8030	This work
LO11153	AN9216 :: <i>AfriboB</i> in LO9721	This work
LO11156	<i>alnD ::AfriboB</i> in LO9721	This work
LO11160	<i>alnE ::AfriboB</i> in LO9721	This work
LO11164	AN9214 :: <i>AfriboB</i> in LO9721	This work
LO11168	AN12145 :: <i>AfriboB</i> in LO9721	This work
LO11178	<i>alnA ::AfriboB</i> in LO9721	This work
LO11181	<i>alnB ::AfriboB</i> in LO9721	This work
LO11184	AN9213 :: <i>AfriboB</i> in LO9721	This work
LO11190	AN9215 :: <i>AfriboB</i> in LO9721	This work
LO11192	<i>alnF ::AfriboB</i> in LO9721	This work
LO11261	<i>alnC ::AfriboB</i> in LO9721	This work
LO11289	<i>alnH ::AfriboB</i> in LO9721	This work
LO11293	AN9221(<i>alnG-alnR</i>) :: <i>AfriboB</i> in LO9721	This work
LO11302	<i>alnI ::AfriboB</i> in LO9721	This work

* An ATG start codon was inserted; AN9221(250-652) was reannotated as A1hR(5-407)

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2.

Hybrid transcription factor induction of the (+)-asperlin gene cluster.

Gene	Hybrid TF (LO9721) ^d		P value ^c (induced vs control)
	Control (LO1362) ^{a,b}	induced	
AN1029 (<i>afbA</i>)	0.30 ± 0.02	891.77 ± 67.24	2.17E-25
AN9223	0.29 ± 0.03	0.26 ± 0.06	--
AN12146	0.29 ± 0.03	0.25 ± 0.06	--
AN11200 (<i>alnI</i>)	2.30 ± 0.51	18.95 ± 1.74	5.19E-10
AN11192 (<i>alnH</i>)	0.29 ± 0.03	743.55 ± 118.12	2.33E-36
AN9221 ^d (<i>alnG</i> and <i>alnR</i>)	14.72 ± 5.69	576.91 ± 83.50	0
AN9220 (<i>alnF</i>)	0.29 ± 0.03	1384.18 ± 194.59	1.17E-40
AN9219 (<i>alnE</i>)	0.29 ± 0.03	1589.13 ± 150.41	2.86E-39
AN9218 (<i>alnD</i>)	0.29 ± 0.03	47.56 ± 12.70	2.60E-28
AN11198 (<i>alnC</i>)	0.29 ± 0.03	38.71 ± 8.74	1.69E-22
AN11199 (<i>alnB</i>)	0.29 ± 0.03	166.57 ± 37.65	4.75E-35
AN11191 (<i>alnA</i>)	0.29 ± 0.03	49.09 ± 15.45	1.13E-33
AN9216	3.57 ± 3.42	5.07 ± 2.11	--
AN9215	0.29 ± 0.03	0.25 ± 0.06	--

^aValues are mean ± standard deviation of reads per million for three replicates. The control strain LO1362 and LO9721, the strain carrying the hybrid transcription factor, were grown in lactose minimal medium which is non-repressing for the *alcA* promoter. After 36 hours of growth, 50 mM MEK was added to half the flasks containing LO9721 to induce expression of the hybrid transcription factor. The remaining LO9721 flasks served as a non-repressing control. MEK was added in parallel to the control strain, LO1362. The hybrid transcription factor drives expression of AN11200-AN11191 (but not other genes in this region) even in non-repressing medium.

^bSeveral of the genes have the same value in the control strain because of a minor quirk of the SeqMonk analysis package. The package normalizes reads per million as base 2 logarithms and 0 cannot be expressed with complete accuracy as a base 2 log (i.e. $0 = 2^{-\infty}$). Conversion from log base 2 values back to numbers results in small fractions of a read in instances in which there were no raw reads. For the Control lane, instances in which there were no raw reads resulted in normalized reads of 0.29.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

p values in the column at the right reveal that the increase in expression in the induced strain vs the control strain is highly significant. P values were obtained using total reads with the R statistical package.

AN9221 is incorrectly annotated in the available databases and is actually two genes that we have designated *alnG* and *alnR*. Induction with MEK results in a substantial additional increase in expression.