



Published in final edited form as:

*Angew Chem Int Ed Engl.* 2016 January 26; 55(5): 1662–1665. doi:10.1002/anie.201507097.

## Development of Genetic Dereplication Strains in *Aspergillus nidulans* Results in the Discovery of Aspercryptin\*\*

**Dr. Yi-Ming Chiang,**

Department of Pharmacology and Pharmaceutical Sciences, School of Pharmacy, University of Southern California, Los Angeles, California 90089 (USA). Department of Pharmacy, Chia Nan University of Pharmacy and Science, Tainan, 71710 (Taiwan)

**Dr. Manmeet Ahuja<sup>§</sup>,**

Department of Molecular Biosciences, University of Kansas, Lawrence, Kansas 66045 (USA)

**C. Elizabeth Oakley,**

Department of Molecular Biosciences, University of Kansas, Lawrence, Kansas 66045 (USA)

**Dr. Ruth Entwistle,**

Department of Molecular Biosciences, University of Kansas, Lawrence, Kansas 66045 (USA)

**Dr. Christoph Zutz,**

Institute for Milk Hygiene, University of Veterinary Medicine Vienna, Vienna, 1210 (Austria)

**Dr. Clay C. C. Wang, and**

Department of Pharmacology and Pharmaceutical Sciences, School of Pharmacy, University of Southern California, Los Angeles, California 90089 (USA). Department of Chemistry, University of Southern California, Los Angeles, California 90089 (USA)

**Dr. Berl R. Oakley**

Department of Molecular Biosciences, University of Kansas, Lawrence, Kansas 66045 (USA)

### Abstract

To reduce the secondary metabolite background in *Aspergillus nidulans* and minimize the rediscovery of compounds and pathway intermediates, we have created a “genetic dereplication” strain in which we have deleted eight of the most highly expressed secondary metabolite gene clusters (more than 244,000 base pairs deleted in total). This strain has allowed us to discover a novel compound that we designate aspercryptin and to propose a biosynthetic pathway for the compound. Interestingly, aspercryptin is formed from compounds produced by two separate gene clusters, one of which makes the well-known product cichorine. This raises the exciting possibility

\*\*This work was supported by the H. L. Snyder Medical Foundation, the Irving S. Johnson Fund of the University of Kansas Endowment and by grant P01-GM084077 from the National Institute of General Medical Sciences. We thank Dr. Anbanandam Asokan and the Bio-NMR Core Facility of the Del Shankel Structural Biology Center at the University of Kansas for 800 MHz NMR data and assistance in solving the structure of aspercryptin.

Correspondence to: Berl R. Oakley.

<sup>§</sup>Current address: Industrial Biotechnology Division, Reliance Technology Group, Reliance Industries Limited, Reliance Corporate Park, Thane Belapur Road, Ghansoli, Navi Mumbai 400701 (India)

Supporting information for this article is given via a link at the end of the document.((Please delete this text if not appropriate))

that fungi use differential regulation of expression of secondary metabolite gene clusters to increase the diversity of metabolites they produce.

## Keywords

genetic dereplication; *Aspergillus*; biosynthesis; NRPS; natural products

Genetic and molecular genetic approaches that up-regulate secondary metabolite (SM) production are dramatically facilitating the discovery of new fungal natural products,<sup>[1]</sup> but these approaches often result in the production of a large number of compounds, including pathway intermediates, and, consequently, complex metabolite profiles. We have thus devised a strategy we call “genetic dereplication” whereby we simplify the discovery of novel compounds by eliminating the major known SM biosynthetic pathways in *A. nidulans*, reducing the complexity of SM profiles such that novel compounds are more easily detected. Elimination of highly expressed biosynthetic pathways might also reserve pools of SM precursors such as acetyl-CoA and malonyl-CoA for pathways expressed at low levels.

Our labs have previously shown that we can delete entire SM biosynthetic gene clusters in *A. nidulans* while recycling a selectable marker.<sup>[2]</sup> We therefore set out to delete, sequentially, as many of the major SM clusters as possible. In this study, we have engineered a strain (LO8030, genotype in Table S1) in which the clusters responsible for the biosynthesis of the following major SM producing clusters are deleted: sterigmatocystin,<sup>[3]</sup> the emericellamides,<sup>[4]</sup> asperfuranone,<sup>[5]</sup> monodictyphenone,<sup>[6]</sup> terrequinone,<sup>[7]</sup> F9775A and B,<sup>[8]</sup> asperthecin,<sup>[9]</sup> and both portions of the split cluster that makes austinol and dehydroaustinol.<sup>[10]</sup> Deletion of these clusters reduced the size of the *A. nidulans* genome by 244,061 bp. LO8030 is surprisingly healthy (Figure 1). It forms aerial hyphae upon prolonged incubation, but it is not difficult to work with nor to transform. As anticipated, production of the products of the deleted SM clusters was eliminated under all conditions resulting in a low SM background.

On glucose minimal medium (GMM) plates, sterigmatocystin (**1**) and the minor SMs terrequinone (**2**) and emericellamides (**3** – **7**) were produced in a parental strain but were eliminated in LO8030 (Figure 2A, trace i – iv). However, one MS (mass spectrometry) detectable peak (**8**, MW = 933) was identified specifically in LO8030 albeit at low yield (Figure 2A, trace v and vi). The molecular formula of **8** was predicted to be C<sub>47</sub>H<sub>79</sub>N<sub>7</sub>O<sub>12</sub> based on its HRESIMS data, suggesting that **8** could be a metabolite biosynthesized by a non-ribosomal peptide synthetase (NRPS) pathway. The MS/MS fragment data of **8** indicated it contains several non-essential amino acids (see below). This is a common feature of metabolites produced from NRPS pathways since these mega-enzymes can incorporate unusual amino acids during the elongation steps. We were able to obtain 3.0 mg of **8** from a large-scale culture. Because of the large size and relatively low solubility of the compound, we took advantage of the high resolution and sensitivity of an 800 MHz spectrometer equipped with a cryoprobe to obtain NMR spectral data. From <sup>13</sup>C-<sup>1</sup>H HSQC-TOCSY (Heteronuclear Single Quantum Coherence-Total Correlation Spectroscopy), six possible spin systems including threonine, isoleucine, aspartic acid/asparagine, serine,

lysine-like, and an unknown spin system could be identified (Figure S1). Analysis of the  $^1\text{H}$ ,  $^{13}\text{C}$ , 2D-NMR data (Figure S2, S7–S9, and Table S2) and MS/MS fragment data (Figure S3) revealed that compound **8** is a cichorine derived hexapeptide with 2-aminododecanol and 2-aminocaprylic acid residues as shown in Figure 2B (for details of structure elucidation, see Supplemental Information). We named compound **8** aspercryptin.

The structure of aspercryptin indicates that it is an NRPS product. There is only one NRPS gene, AN7884, in the genome of *A. nidulans* containing six adenylate (A) domains.<sup>[11]</sup> Andersen et al.,<sup>[12]</sup> found that the genes AN7872 to AN7884 are co-regulated, suggesting that they may form an SM gene cluster. Bioinformatic analyses of genes in this cluster indicate that they encode a fatty acid synthase, aminotransferases, a P450 hydroxylase, a short chain dehydrogenase, and transporters, most of them reasonably predicted to be involved in the biosynthesis of aspercryptin (Figure 3).

To determine if this gene cluster encodes the aspercryptin biosynthetic pathway, we deleted the NRPS gene AN7884 and, indeed, deletion of AN7884 eliminated the production of aspercryptin (Figure 2A, trace vii), indicating that aspercryptin is synthesized by the AN7884 cluster. We designate AN7884 as *atnA* and give the other genes of the cluster the designations shown in Figure 3. Aspercryptin was found to be produced at a higher titer when LO8030 was grown on YAG (yeast extract, agar, and glucose) plates. In this condition, a new aspercryptin derivative **9** was identified, albeit at lower intensity than aspercryptin (Figure S4A). The MW of compound **9** is 14 Da less than aspercryptin and the MS/MS fragment data indicated that **9** is aspercryptin with Ile replaced by Val (Figure S5). To gain insights into the functions of *atn* genes, we deleted 12 additional genes in the gene cluster (from *atnB* to *atnM*) and analyzed the metabolites produced on YAG plates (Figure S4). Deletion of the two fatty acid synthase genes, *atnF* and *atnM*, eliminated > 99.5% of **8** indicating that the two FAS subunit genes are necessary for aspercryptin biosynthesis. The tiny amount of **8** detected in the FAS deletion strains suggests that fatty acids generated by other, endogenous FASs can be incorporated, albeit at a significantly lower rate. Deletion of the dehydrogenase (*atnD*) and cytochrome P450 (*atnE*) genes abolished **8**. AtnD and AtnE are likely involved in the oxidation of the  $\alpha$  carbon of the fatty acids. Deletion of the amino acid aminotransferase genes, *atnJ* and *atnH*, eliminated approximately ~70% and ~40% of **8**, respectively (Figure S4B), suggesting that AtnJ and AtnH might be able to compensate for each other but with less efficiency, or that other endogenous aminotransferases could partially compensate for their functions. Deletion of *atnI*, an RTA1 superfamily protein, eliminated > 70% of **8**. Deletion of three genes, an MFS transporter (*atnC*), an ABC transporter (*atnG*), and a thioesterase-like gene (*atnL*) did not alter the yield of **8** significantly. Interestingly, deletion of *atnB* and *atnK*, decreased the titer of **8** by greater than 80% but increased the titer of **9** more than six fold, suggesting that AtnB and AtnK might participate in the selection of Ile versus Val residue in module 3 of AtnA. Without AtnB or AtnK, the module 3 of AtnA prefers to incorporate Val and produces **9**. Analysis of the deletion mutants allowed us to propose a biosynthetic pathway for aspercryptin (Figure 4). The fatty acid synthase subunits (AtnF and AtnM), dehydrogenase (AtnD) and cytochrome P450 (AtnE) and amino acid transferases (AtnJ and AtnH) could be involved in the biosynthesis of two unusual amino acids, 2-aminocaprylic and 2-aminododecanoic acids,

which are then activated and incorporated into the growing peptide chain by AtnA. Inspection of the domain architecture of AtnA revealed that the second module has an epimerase (E) domain, suggesting the second residue of **8** could be D-allo-Thr. Indeed, Marfey's analysis<sup>[13]</sup> showed that **8** does contain D-allo-Thr and the L form of the remaining amino acids (Figure S6). After condensation of the hexapeptide of **8**, the C-terminal reductase (R) domain might be involved in the reductive release and production of the aldehyde hexapeptide.<sup>[14]</sup> Further reduction would generate **8**.

Our proposed pathway postulates that aspercryptin synthesis uses cichorine-Ser as a precursor. Cichorine is a phytotoxin originally discovered as a natural product produced by members of the genus *Alternaria*. We have previously identified the cichorine biosynthetic cluster in *A. nidulans* and found that the non-reducing polyketide synthase AN6448 is essential for cichorine biosynthesis.<sup>[15]</sup> If cichorine-Ser produced by the cichorine pathway is the precursor of aspercryptin, deletion of AN6448 should eliminate aspercryptin production. We, consequently, deleted AN6448 and found that production of aspercryptin is, indeed, eliminated (Figure 2A, trace viii).

Our data demonstrate that genetic dereplication strains are valuable in discovering novel compounds from unknown biosynthesis pathways. The fact that the SM background is lowered dramatically may be particularly valuable in detecting compounds produced by heterologous expression of SM genes from other fungi and in searches for genes that regulate cryptic SM clusters. They may have an additional advantage in production of compounds or proteins for human and animal use because they are incapable of producing major toxic SMs such as sterigmatocystin (**1**). Our data also lead us to the fascinating conclusion that aspercryptin is made of building blocks from two distinct clusters that are physically separated in the genome, the AN6448 (cichorine) cluster and the *atn* cluster. They raise the interesting possibility that differential gene regulation could result in the production of cichorine or aspercryptin. This, in turn, raises the exciting possibility that *A. nidulans* (and by inference other fungi) may use differential regulation of SM gene cluster expression to expand their repertoire of natural products and tailor their SM arsenal to achieve maximum competitive advantage. Forseth et al. found that a somewhat similar situation occurs with the *Ina* and *Inb* clusters in *Aspergillus flavus*.<sup>[16]</sup> These clusters are closely related to each other and contain homologous NRPS-like core biosynthetic genes. They are partially redundant in that when the *Ina* NRPS-like gene (*InaA*) is deleted, the *Inb* NRPS-like gene (*InbA*) and perhaps other genes in the *Inb* cluster can modify intermediates produced by the *Ina* cluster to produce two of the major products of the *Ina* cluster, albeit in greatly reduced amounts. In addition, Forseth et al. also found evidence that intermediates of the *Ina* biosynthetic pathway are modified by the *Inb* biosynthetic pathway. The situation with the cichorine and aspercryptin clusters differs from the *Ina* and *Inb* situation in significant ways, however. The cichorine and aspercryptin clusters are not homologous and the core biosynthetic enzyme for the cichorine pathway is a PKS whereas the core biosynthetic gene for the aspercryptin cluster is an NRPS. The cichorine and aspercryptin pathways are not redundant. Rather they function in a serial fashion with the cichorine cluster able to produce an important metabolite on its own, and the *atn* cluster using that product to produce aspercryptin.

## Experimental Section

Deletion of entire SM clusters was carried out as previously described.<sup>[2]</sup> Most clusters were deleted using the loop out recombination procedure based on the procedure of Takahashi *et al.*<sup>[17]</sup> Correct deletion of clusters was verified by diagnostic PCR amplifications using primers outside of the ends of the clusters. *Atn* cluster genes were deleted and deletions were verified by diagnostic PCR using the methods of Oakley *et al.*<sup>[18]</sup> Growth media and conditions, as well as extraction conditions, are detailed in the supplemental materials.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## References

1. Keller NP, Turner G, Bennett JW. *Nat Rev Microbiol.* 2005; 3:937. [PubMed: 16322742] Bayram O, Braus GH. *FEMS Microbiol Rev.* 2012; 36:1. [PubMed: 21658084] Brakhage AA. *Nat Rev Microbiol.* 2013; 11:21. [PubMed: 23178386] Yu JH, Keller N. *Annu Rev Phytopathol.* 2005; 43:437. [PubMed: 16078891] Calvo AM. *Fungal Genet Biol.* 2008; 45:1053. [PubMed: 18457967] Hoffmeister D, Keller NP. *Nat Prod Rep.* 2007; 24:393. [PubMed: 17390002] Fox EM, Howlett BJ. *Curr Opin Microbiol.* 2008; 11:481. [PubMed: 18973828] Hertweck C. *Nat Chem Biol.* 2009; 5:450. [PubMed: 19536102] Georgianna DR, Fedorova ND, Burroughs JL, Dolezal AL, Bok JW, Horowitz-Brown S, Woloshuk CP, Yu J, Keller NP, Payne GA. *Mol Plant Pathol.* 2010; 11:213. [PubMed: 20447271] Yin W, Keller NP. *J Microbiol.* 2011; 49:329. [PubMed: 21717315] Sanchez JF, Somoza AD, Keller NP, Wang CC. *Nat Prod Rep.* 2012; 29:351. [PubMed: 22228366]
2. Chiang YM, Oakley CE, Ahuja M, Entwistle R, Schultz A, Chang SL, Sung CT, Wang CCC, Oakley BR. *J Am Chem Soc.* 2013; 135:7720. [PubMed: 23621425]
3. Brown DW, Adams TH, Keller NP. *Proc Natl Acad Sci U S A.* 1996; 93:14873. [PubMed: 8962148]
4. Chiang YM, et al. *Chem Biol.* 2008; 15:527. [PubMed: 18559263]
5. Chiang YM, Szcwcyk E, Davidson AD, Keller N, Oakley BR, Wang CCC. *J Am Chem Soc.* 2009; 131:2965. [PubMed: 19199437]
6. Chiang YM, Szcwcyk E, Davidson AD, Entwistle R, Keller NP, Wang CCC, Oakley BR. *Appl Environ Microbiol.* 2010; 76:2067. [PubMed: 20139316]
7. Bok JW, Hoffmeister D, Maggio-Hall LA, Murillo R, Glasner JD, Keller NP. *Chem Biol.* 2006; 13:31–37. [PubMed: 16426969]
8. Sanchez JF, Chiang YM, Szcwcyk E, Davidson AD, Ahuja M, Oakley CE, Bok JW, Keller N, Oakley BR, Wang CCC. *Mol Biosyst.* 2010; 6:587. [PubMed: 20174687]
9. Szcwcyk E, Chiang YM, Oakley CE, Davidson AD, Wang CCC, Oakley BR. *Appl Environ Microbiol.* 2008; 74:7607. [PubMed: 18978088]
10. Lo HC, Entwistle R, Guo CJ, Ahuja M, Szcwcyk E, Hung JH, Chiang YM, Oakley BR, Wang CC. *J Am Chem Soc.* 2012; 134:4709. [PubMed: 22329759]
11. von Dohren H. *Fungal Genet Biol.* 2009; 46(Suppl 1):S45. [PubMed: 18804170]
12. Andersen MR, et al. *Proc Natl Acad Sci U S A.* 2013; 110:E99. [PubMed: 23248299]
13. Marfey P. *Carlberg Res Commun.* 1984; 49:591.
14. Gaitatzis N, Kunze B, Müller R. *Proc Natl Acad Sci U S A.* 2001; 98:11136. [PubMed: 11562468] Du L, Lou L. *Nat Prod Rep.* 2010; 27:255. [PubMed: 20111804]
15. Sanchez JF, Entwistle R, Corcoran D, Oakley BR, Wang CCC. *Med Chem Commun.* 2012; 3:997.
16. Forseth RR, Amaike S, Schwenk D, Affeldt KJ, Hoffmeister D, Schroeder FC, Keller NP. *Angew Chem Int Ed Engl.* 2013; 52:1590. [PubMed: 23281040]
17. Takahashi T, Jin FJ, Sunagawa M, Machida M, Koyama Y. *Appl Environ Microbiol.* 2008; 74:7684. [PubMed: 18952883]

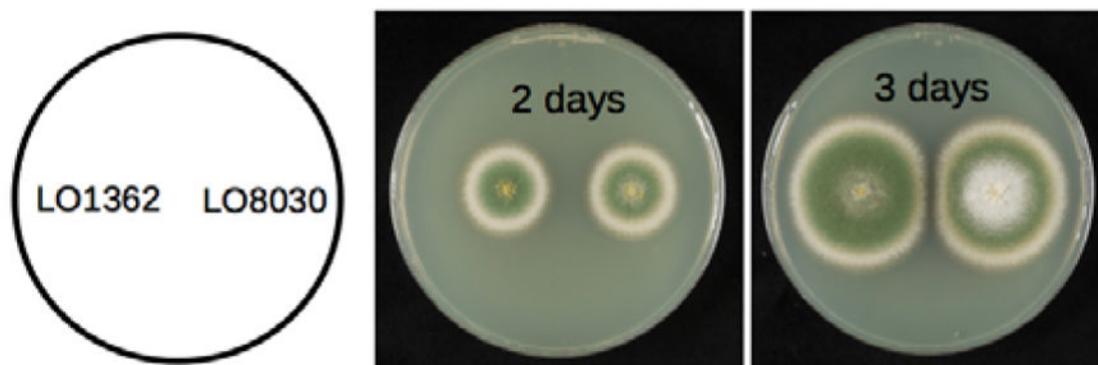
18. Oakley CE, Edgerton-Morgan H, Oakley BR. *Methods Mol Biol.* 2012; 944:143. [PubMed: 23065614]

Author Manuscript

Author Manuscript

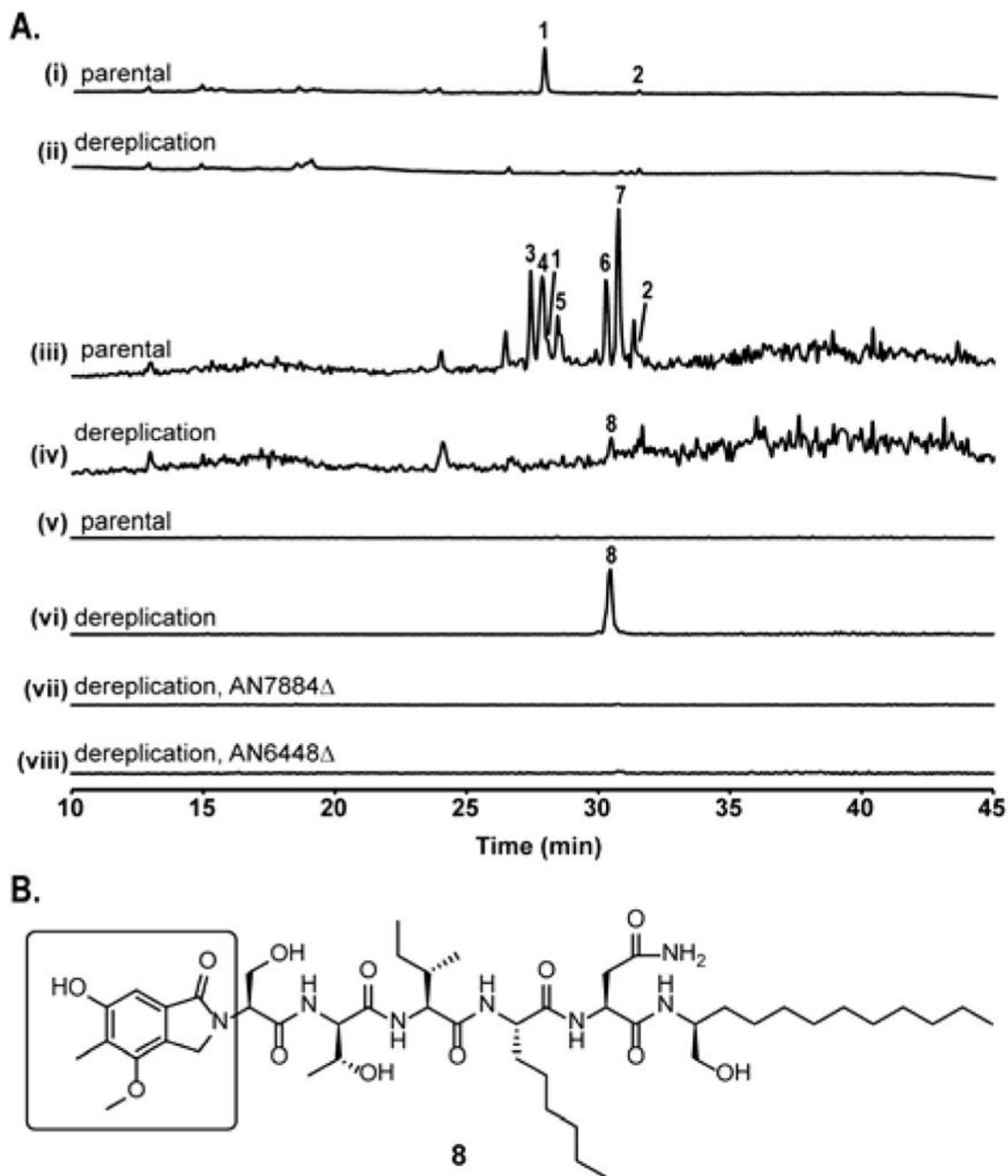
Author Manuscript

Author Manuscript



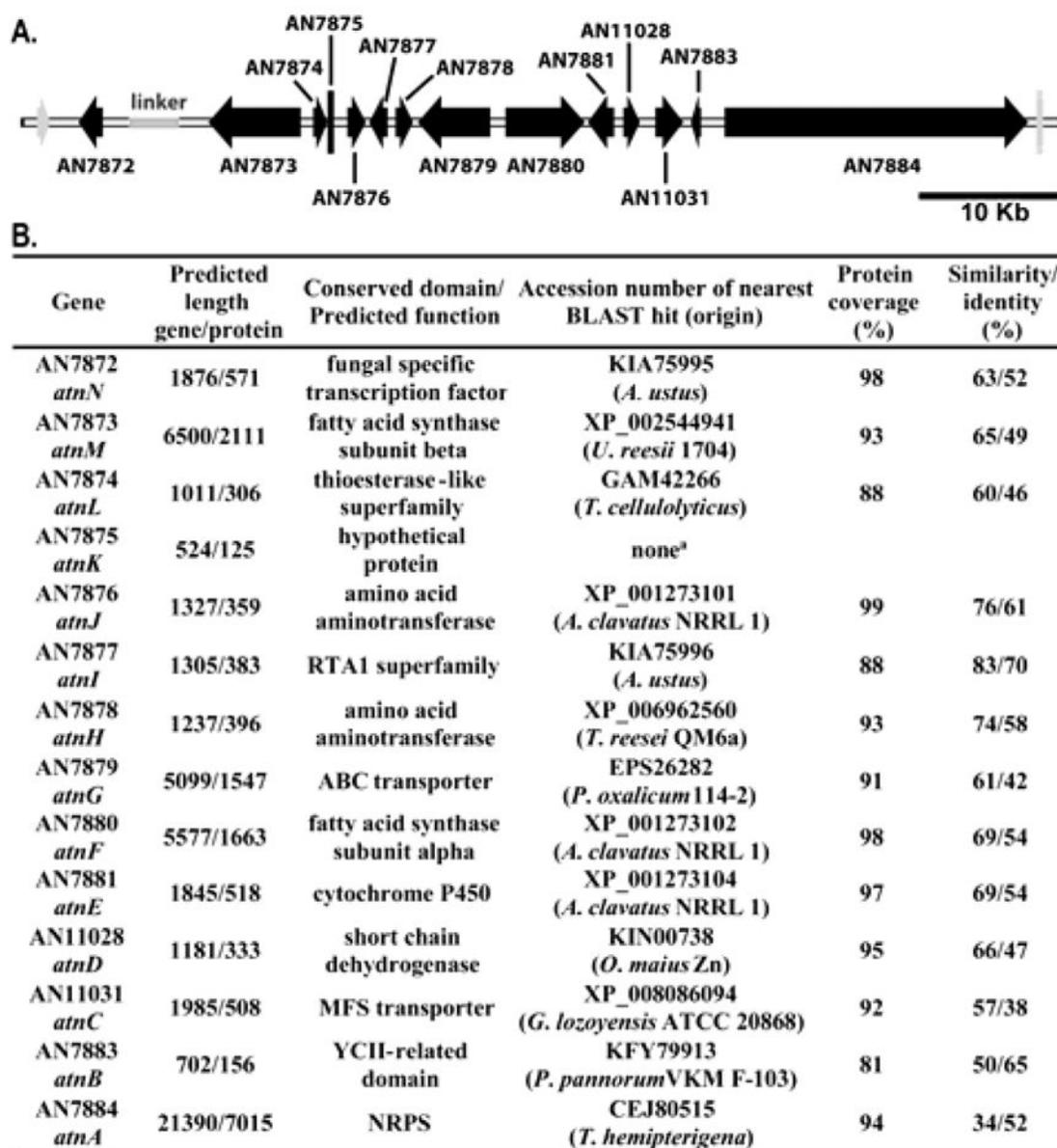
**Figure 1.**

Growth of a parental strain, LO1362, and a daughter strain, LO8030, in which eight secondary metabolism gene clusters have been deleted. Growth is on complete medium. After two days, the multi-cluster deletion strain grows as well as its parental strain. After three days, radial growth for the multi-cluster deletion strain is the same as for the parent. However, an increase in aerial hyphae makes the center of the colony white.



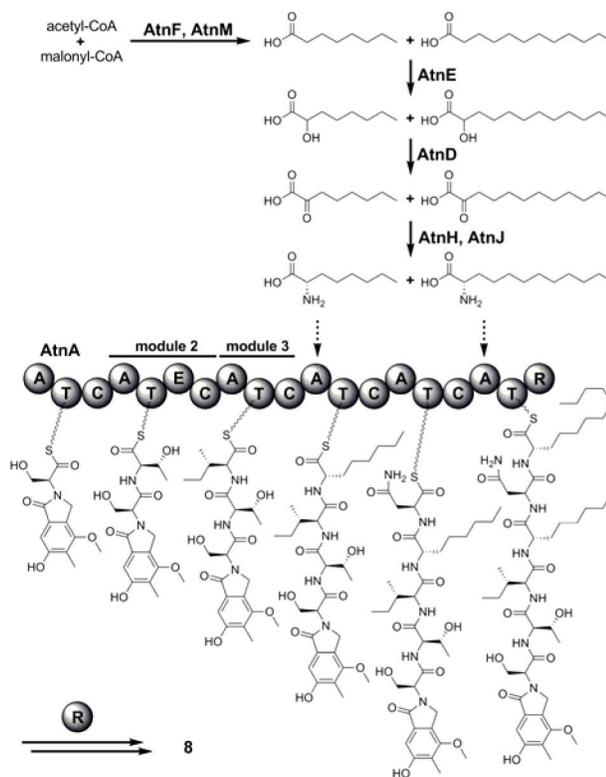
**Figure 2.**

**A.** HPLC metabolic profiles of the parental and genetic dereplication strains. UV-Vis total scan from 200 – 600 nm (i, ii), total ion current (iii, iv), and extracted ion chromatogram at  $m/z$  934 (v – viii). All strains were grown on GMM plates. **1** is sterigmatocystin, **2** is terrequinone, **3 – 7** are emericellamide C, D, A, E, and F, respectively, **8** is aspercryptin. **B.** The chemical structure of aspercryptin (**8**). The box indicates the cichorine moiety. Strains: parental = LO1362; dereplication = LO8030; dereplication, AN7884 = LO8096; dereplication, AN6448 = LO9345. Genotypes are given in Table S1.



**Figure 3.**

**A.** Schematic representation of the AN7884 (*atn*) cluster. Each arrow indicates the direction of transcription and relative sizes of the open reading frames (ORFs). This cluster spanned contig 134 and 135. The linker represents the connection between contig 134 and 135. **B.** The deduced function of each ORF and the amino acid sequence coverage and similarity/identity, as compared with the BLAST search of the NCBI nonredundant protein database. Genetic symbols we are assigning to members of the aspercryptin cluster are below the AspGD gene numbers in column 1. <sup>a</sup>The nearest hit to AN7875 only has 40% of protein coverage and was considered to be irrelevant.



**Figure 4.**  
Proposed biosynthesis pathway of aspercryptin (**8**).