



Acurin A, a novel hybrid compound, biosynthesized by individually translated PKS- and NRPS-encoding genes in *Aspergillus aculeatus*

Wolff, Peter B.; Nielsen, Maria L.; Slot, Jason C.; Andersen, Lasse N.; Petersen, Lene M.; Isbrandt, Thomas; Holm, Dorte K.; Mortensen, Uffe H.; Nødvig, Christina S.; Larsen, Thomas O.

Total number of authors:

11

Published in:

Fungal Genetics and Biology

Link to article, DOI:

[10.1016/j.fgb.2020.103378](https://doi.org/10.1016/j.fgb.2020.103378)

Publication date:

2020

Document Version

Peer reviewed version

[Link back to DTU Orbit](#)

Citation (APA):

Wolff, P. B., Nielsen, M. L., Slot, J. C., Andersen, L. N., Petersen, L. M., Isbrandt, T., Holm, D. K., Mortensen, U. H., Nødvig, C. S., Larsen, T. O., & Hoof, J. B. (2020). Acurin A, a novel hybrid compound, biosynthesized by individually translated PKS- and NRPS-encoding genes in *Aspergillus aculeatus*. *Fungal Genetics and Biology*, 139, Article 103378. <https://doi.org/10.1016/j.fgb.2020.103378>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Journal Pre-proofs

Acurin A, a novel hybrid compound, biosynthesized by individually translated PKS- and NRPS-encoding genes in *Aspergillus aculeatus*

Peter B. Wolff, Maria L. Nielsen, Jason C. Slot, Lasse N. Andersen, Lene M. Petersen, Thomas Isbrandt, Dorte K. Holm, Uffe H. Mortensen, Christina S. Nødvig, Thomas O. Larsen, Jakob B. Hoof

PII: S1087-1845(20)30069-4
DOI: <https://doi.org/10.1016/j.fgb.2020.103378>
Reference: YFGBI 103378

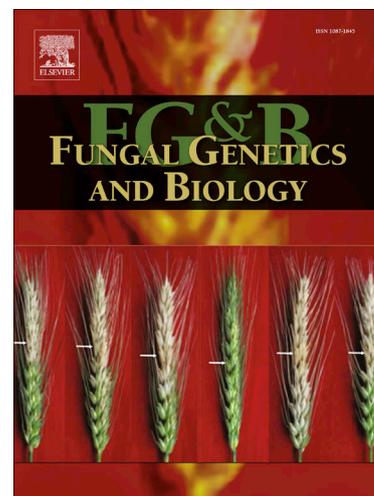
To appear in: *Fungal Genetics and Biology*

Received Date: 6 January 2020
Revised Date: 10 March 2020
Accepted Date: 12 March 2020

Please cite this article as: Wolff, P.B., Nielsen, M.L., Slot, J.C., Andersen, L.N., Petersen, L.M., Isbrandt, T., Holm, D.K., Mortensen, U.H., Nødvig, C.S., Larsen, T.O., Hoof, J.B., Acurin A, a novel hybrid compound, biosynthesized by individually translated PKS- and NRPS-encoding genes in *Aspergillus aculeatus*, *Fungal Genetics and Biology* (2020), doi: <https://doi.org/10.1016/j.fgb.2020.103378>

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2020 Published by Elsevier Inc.



Acurin A, a novel hybrid compound, biosynthesized by individually translated PKS- and NRPS-encoding genes in *Aspergillus aculeatus*

Peter B. Wolff^{a,1}, Maria L. Nielsen^{a,1}, Jason C. Slot^b, Lasse N. Andersen^a, Lene M. Petersen^a, Thomas Isbrandt^a, Dorte K. Holm^a, Uffe H. Mortensen^a, Christina S. Nødvig^a, Thomas O. Larsen^{*a}, Jakob B. Hoof^{*a}

^a Department of Biotechnology and Biomedicine (DTU Bioengineering), Technical University of Denmark, Søtofts Plads, Building 221/223, 2800 Kgs. Lyngby (Denmark). *E-mail: tol@bio.dtu.dk/jblni@dtu.dk.

^b Department of Plant Pathology, The Ohio State University, 481C Kottman Hall, 2021 Coffey Road, Columbus, OH 43210 (USA).

¹ These authors have contributed equally to this work.

Highlights

- A biosynthetic gene cluster for the hybrid polyketide-nonribosomal peptide compound acurin A in *Aspergillus aculeatus*.
- The PKS and NRPS responsible for acurin A production are two individually transcribed entities.
- Phylogenetics suggest that the separated PKS and NRPS is a result from the fission of an ancestral PKS-NRPS hybrid.
- The biosynthetic capacity of *A. aculeatus* is facilitated by the use of CRISPR/Cas9 gene editing.

Acurin A, a novel hybrid compound, biosynthesized by individually translated PKS- and NRPS-encoding genes in *Aspergillus aculeatus*

Peter B. Wolff^{a,1}, Maria L. Nielsen^{a,1}, Jason C. Slot^b, Lasse N. Andersen^a, Lene M. Petersen^a, Thomas Isbrandt^a, Dorte K. Holm^a, Uffe H. Mortensen^a, Christina S. Nødvig^a, Thomas O. Larsen^{*a}, Jakob B. Hoof^{*a}

^a Department of Biotechnology and Biomedicine (DTU Bioengineering), Technical University of Denmark, Søtofts Plads, Building 221/223, 2800 Kgs. Lyngby (Denmark). *E-mail: tol@bio.dtu.dk/jblni@dtu.dk.

^b Department of Plant Pathology, The Ohio State University, 481C Kottman Hall, 2021 Coffey Road, Columbus, OH 43210 (USA).

¹ These authors have contributed equally to this work.

Abstract

This work presents the identification and proposed biosynthetic pathway for a compound of mixed polyketide-nonribosomal peptide origin that we named acurin A. The compound was isolated from an extract of the filamentous fungus *Aspergillus aculeatus*, and its core structure resemble that of the mycotoxin fusarin C produced by several *Fusarium* species. Based on bioinformatics in combination with RT-qPCR experiments and gene-deletion analysis, we identified a biosynthetic gene cluster (BGC) in *A. aculeatus* responsible for the biosynthesis of acurin A. Moreover, we were able to show that a polyketide synthase (PKS) and a nonribosomal peptide synthetase (NRPS) enzyme separately encoded by this BGC are responsible for the synthesis of the PK-NRP compound, acurin A, core structure. In comparison, the production of fusarin C is reported to be facilitated by a linked PKS-NRPS hybrid enzyme. Phylogenetic analyses suggest the PKS and NRPS in *A. aculeatus* resulted from a recent fission of an ancestral hybrid enzyme followed by gene duplication. In addition to the PKS- and NRPS-encoding genes of acurin A, we show that six other genes are influencing the biosynthesis including a regulatory transcription factor. Altogether, we have demonstrated the involvement of eight genes in the biosynthesis of acurin A, including an in-cluster transcription factor. This study highlights the biosynthetic capacity of *A. aculeatus* and serves as an example of how the CRISPR/Cas9 system can be exploited for the construction of fungal strains that can be readily engineered.

Keywords: *Aspergillus aculeatus*, CRISPR/Cas9, Secondary metabolism, Acurin A, hybrid compound, PKS, NRPS.

1. Introduction

The *Aspergillus* genus contains several species that are used as industrial workhorses in biotechnology. Importantly, this includes species that are potential producers of mycotoxins that may pose a risk to human health (Bennett and Klich, 2003). For example, *Aspergillus niger* is a recognized producer of both ochratoxin A and fumonisin B2 (Frisvad et al., 2011). Importantly, by identifying the genetics that is necessary for the production of mycotoxins, it is possible to create safe production strains as the relevant secondary metabolite (SM) encoding genes can be eliminated. *Aspergillus aculeatus* is known to produce several industrially relevant enzymes (Banerjee et al., 2007; Suwannarangsee et al., 2014) and numerous bioactive SMs (Andersen et al., 1977; Hayashi et al., 1999; Petersen et al., 2014), e.g., secalonic acid toxins, the anti-insectan okaramines, and the antifungal calbistrins. Moreover, an analysis of its genome has uncovered 71 Biosynthetic Gene Clusters (BGCs), indicating that its potential for the production of SMs is substantial (“Secondary Metabolism Clusters - *Aspergillus aculeatus* ATCC16872 v1.1,” n.d.; Vesth et al., 2018). The BGC count are based on a locus, which include a gene that encode the main scaffold of the SM; either polyketide synthases (PKSs, 23), non-ribosomal synthetases (NRPSs, 18), PKS-NRPS hybrids (4), NRPS-likes (21), terpene cyclases (TCs, 4)/prenyl transferases (PTs, 3). The scaffold producing enzymes utilizes different building blocks. Thus, PKSs typically polymerize acyl-CoA starter units with malonyl-CoA extender units for synthesis, NRPSs require ATP-activated amino acids (proteinogenic and non-proteinogenic), whereas TC/PT uses C5-isoprenoid units. Depending on the respective enzyme type and organization, the PKS may, or may not, reduce and methylate the scaffold, where the specificity and promiscuity of amino-acid selection in NRPSs reside in the adenylation domain. In addition, the genes that code for the scaffold function are typically flanked by other genes that participate in the biosynthesis of the SM by decorating or modifying the scaffold(s) or regulates the production. Assigning genes to a BGC requires experimental validation.

To date, only a few SMs (e.g., aculinic acid, aculenes, and Acu-dioxomorpholine A (Lee et al., 2019; Petersen et al., 2015; Robey et al., 2018)) have been linked experimentally to their corresponding BGC. For example, aculinic acid is synthesized from the polyketide (PK) 6-methylsalicylic acid (6-MSA), which is a precursor of other PKs, including the mycotoxin patulin (Puel et al., 2010). Therefore, considering the industrial potential of *A. aculeatus*, it is

highly desirable to improve the genetic map of its SM metabolism to identify new bioactive compounds and uncover potential mycotoxins (Clevenger et al., 2017).

Recently, efficient genetic tools, including CRISPR/Cas technology, have been introduced to *A. aculeatus*. To set the stage for efficient SM gene discovery, we have further elaborated on the genetic toolbox for *A. aculeatus* by constructing a non-homologous end joining (NHEJ) deficient strain mediated by CRISPR/Cas9, where subsequent gene-targeting efforts becomes highly efficient. Using bioinformatics, our chemical-analysis platform, and genetic toolbox, we set out to explore an unknown compound with the exact mass 373.1889 Da that we had previously identified in *A. aculeatus* (Petersen et al., 2015). Here we show that this compound is related to the mycotoxin fusarin C, which is produced by several *Fusarium* species (Cantalejo et al., 1999; Gelderblom et al., 1984; Niehaus et al., 2013; Song et al., 2004). We also identified the BGC responsible for the production of acurin A and proposed a biosynthetic pathway for its biosynthesis. Phylogenetic analyses suggest the PKS and NRPS in *A. aculeatus* resulted from a recent fission of an ancestral hybrid enzyme followed by gene duplication.

2. Materials and Methods

2.1. Strains and media

A list of all strains constructed in this study is provided in Table S1. Gene deletions in *A. aculeatus* was performed using strain ACU11 (*pyrG1*, *akuAΔ*), while overproduction of transcription factor AcrR was performed either ectopically for screening purposes using the NHEJ-proficient strain ACU9 (*pyrG1*) (Nødvig et al., 2015), or at a specific locus in ACU11 for gene deletions; both ACU11 and ACU9 are derived from the wild-type (WT) genome sequenced *A. aculeatus* ATCC16872, and have growth requirement for uridine. We used annotation v1.1 of the *A. aculeatus* genome sequence (Joint Genome Institute, JGI) or GenBank, NCBI, to retrieve relevant sequence information. Genomic DNA was extracted from *A. aculeatus* ATCC16872 using the FastDNA™ SPIN Kit for Soil DNA extraction (MP Biomedicals, USA). *Escherichia coli* strain DH5α was used for plasmid propagation. *Aspergillus* solid and liquid minimal medium (MM), transformation medium (TM), and yeast extract sucrose (YES) medium was supplemented with 10 mM uridine and 10 mM uracil when necessary. MM and TM were prepared as described by Nødvig et al. (Nødvig et al., 2015), and YES was prepared as described in Samson et al. (Samson et al., 2010). *E. coli* was grown on

solid- and in liquid Luria-Bertani (LB) medium containing 10 g/l tryptone (Bacto), 5 g/l yeast extract (Bacto), and 10 g/l NaCl (pH 7.0) supplemented with 100 µg/ml ampicillin.

2.2. Vector- and strain construction

All primers (Integrated DNA Technology, Belgium) used for amplification of *A. aculeatus* derived PCR fragments and qPCR are listed in Table S2. PCR fragments were generated using the PfuX7 polymerase (Nørholm, 2010), and plasmids were constructed by Uracil-Specific Excision Reagent (USER) fusion as described previously (Hansen et al., 2011; Nielsen et al., 2013) and were purified using the GenElute™ Plasmid Miniprep Kit (Sigma-Aldrich). All constructed plasmids for gene deletions contained two *SwaI* restriction sites, the ampicillin-resistance gene and an origin of replication for propagation in *E. coli*. Gene-targeting vectors for gene deletions were constructed based on pU2002A by introducing the PCR-amplified up- and downstream fragments into two *PacI*/Nt.BbvCI USER cassettes on each side of either the *A. fumigatus*/*A. flavus pyrG* marker (AF*pyrG* and AFL*pyrG*, respectively) flanked by a direct repeat (DR) sequence for marker recycling (pU2002A, pAC3) (Hansen et al., 2011). As all gene deletions of putative BGC residing genes were carried out based on the *pyrG1 akuAΔ* strain background. The *akuA* (Protein ID 127270, encoding the NHEJ factor Ku70) was deleted in *A. aculeatus pyrG1* by co-transforming a linearized gene-targeting substrate containing AFL*pyrG* gene flanked by a DR and an *argB*/AMA1-based CRISPR/Cas gene-editing vector containing *cas9* and *akuA*-targeting sgRNA, see Table S2 and for experimental description (Nødvig et al., 2015). The gene-expression vector (GEV) for ectopic overexpression of the *acrR* transcription factor in the *A. aculeatus pyrG1* strain was constructed by cloning of the PCR amplified *acrR* fragment into a vector (pU2115-1) containing the constitutive *A. nidulans* promoter *PgpdA*, the terminator *TtrpC*, and AF*pyrG* as well as targeting sequences for integration site 1, *ISI*, in *A. nidulans* (Hansen et al., 2011). Alternatively, the GEV for specific integration into *ISI* in *A. aculeatus* was assembled from pU2002A using a combined upstream targeting sequence with *PgpdA* and *TtrpC* split by a *PacI*/Nt.BbvCI USER cassette as one cloning part together with the downstream targeting sequence. After vector digestion with *PacI* and Nt.BbvCI, this vector was ready to receive *acrA*. The selection of AACU_*ISI* targeting sequences were based on non-coding sequences placed close to what we believed were conserved and highly expressed genes as targeting regions. In this case between AACU_117954 and AACU_59915. Vectors for gene targeting/expression were linearized prior to transformation with *SwaI* (New England Biolabs).

2.3. Protoplastation and transformation

Spores from a single agar plate of *A. aculeatus* were harvested and grown over night in liquid yeast extract peptone dextrose (YPD) medium at 30 °C while shaking. The biomass was filtered through sterile Miracloth (EMD Millipore), washed with sterile Milli-Q water, and subsequently digested in a 50 ml Falcon tube in a 20 ml solution of 40 mg/ml Glucanex (Novozymes A/S) while shaking. For digestion, the biomass was incubated at 30 °C for approximately 3 hours at 150 rpm. Following digestion, the biomass was again filtered through sterile Miracloth and the flow-through was collected and centrifuged at 1200 x g for 10 min. The supernatant was discarded and the pellet was washed twice in ST buffer (1.0 M sorbitol; 50 mM Tris, pH 8.0). The final pellet was resuspended in approximately 2 ml STC buffer (1.0 M sorbitol; 50 mM Tris, pH 8.0; 50 mM CaCl₂) for a final concentration of 1.2x10⁷ cells/ml. PCT buffer (40 % PEG 4000 in STC) was added to the protoplast suspension for a final concentration of approximately 20 % PEG. Transformation was performed using linearized plasmids as described by Nødvig et al. (Nødvig et al., 2015). Transformation plates were incubated at 30 °C until colonies appeared on the plates (5-10 days). All strains were verified as homokaryotic gene-deletion mutants by diagnostic tissue-PCR (see Fig. S1 for validation scheme and Fig. S2-7 for analysis of strains, and Table S3 for analysis parameters).

2.4. Molecular evolution analyses

Homologs of BGCs in a database of 475 fungal genomes (Table S4) were defined as overlapping pairs of query gene homologs (as retrieved by usearch version 8.0.1517 (Edgar, 2010) using the ublast algorithm; sequences with at least 45 % amino acid similarity were retained) with a maximum of six intervening genes in each genome. Hybrid KS (Blin et al., 2019) and AMP binding (PF00501.28) domains were extracted from protein sequences using hmmer (v. 3.1b2) for analysis separately. For each domain and all unique genes in acurin and fusarin C BGC, phylogenies were constructed by first aligning amino acid sequences with mafft (v. 7.221) (Katoh and Standley, 2013) and removing ambiguously aligned characters with Trimal (v. 1.4) (Capella-Gutiérrez et al., 2009) using the automated method. Preliminary phylogenetic analyses constructed by fasttree (v. 2.1.7) (Price et al., 2010) were used to manually reduce alignments to strongly supported bipartitions, where required. Final phylogenies were inferred using maximum likelihood in IQ-TREE (v. 1.6.12) (Nguyen et al., 2015), with standard model selection and branch support assessed by 1000 ultrafast bootstraps.

2.5. RNA purification and RT-qPCR

RNA purification was conducted with the RNeasy Plus Mini Kit (Qiagen) in accordance with the supplied manual with an additional pre-step. A small scrape from a selected colony was transferred to a sterile 2 ml Eppendorf tube together with 1-3 sterile metal beads ($d = 3$ mm). The samples were placed in dry ice or liquid nitrogen for 2-3 min before being transferred to a TissueLyser LT (Qiagen). The samples were lysed at a frequency of 45 Hz for 1.5 min. 350 μ l RLT plus (from the kit) with 1 % β -mercaptoethanol was added and following were as described in manufacturer's protocol. Complementary DNA synthesis was performed using the QuantiTech Reverse Transcription Kit (Qiagen) in accordance with manufacturer's protocol, where all RNA samples were adjusted to 700 ng RNA/20 μ l. RNA concentrations were measured using a Nanodrop lite spectrophotometer (Thermo Scientific). The QuantiTech SYBR Green RT-PCR kit was used as mastermix for analysis of the selected genes. All primers were designed with a T_m range of 58-62 $^{\circ}$ C. The fragments were designed to be between 150-300 bp and, if possible, traversing an intron to distinguish mRNA from gDNA. All samples were analyzed on a Bio-Rad GFX connect Real-Time system. The PCR reaction mixture was as follows: 10 μ l SYBR green, 2 μ l of primer-f (10 M), 2 μ l of primer-r (10 M), 2 μ l diluted cDNA (25 times diluted), and Milli-Q water up to 20 μ l. The PCR program was 95 $^{\circ}$ C for 5 min followed by 40 cycles of 95 $^{\circ}$ C for 10 s and 60 $^{\circ}$ C for 30 s. A melting curve test from 65 $^{\circ}$ C to 95 $^{\circ}$ C with reads every 0.2 min ended the program to evaluate the purity of the reaction products. All RT-qPCR data were normalized to two housekeeping genes; *actA* (Protein ID 1882512) and *hhtA* (Protein ID 1892954) for actin and histone 3, respectively. The normalized data were calculated using the following equations: Fold change = 2^x , $x = (\text{geneOE} - \text{housekeepingOE}) - (\text{generef.} - \text{housekeepingref.})$. All RT-qPCR data were analyzed in triplicates and submitted to a significance test using a two-tailed t-test provided by Microsoft Excel.

2.6. Chemical analysis

All validated strains were cultivated on solid YES and MM medium as three-point stabs and incubated for 7 days at 30 $^{\circ}$ C. Plug extractions were performed as Smedsgaard et al. (Smedsgaard, 1997) with the exception that secondary metabolites were extracted with 3:1 ethylacetate:isopropanol containing 1% formic acid. The samples were analyzed on a maXis

3G orthogonal acceleration quadrupole time-of-flight mass spectrometer (Bruker Daltonics) equipped with an electrospray ionization (ESI) source and connected to an Ultimate 3000 UHPLC system (Dionex), equipped with a Kinetex 2.6 μm C18, 100 mm x 2.1 mm column (Phenomenex). The method applied was described by Holm et al. (Holm et al., 2014). MS/MS analysis were performed on an Agilent Infinity 1290 UHPLC system (Agilent Technologies, Santa Clara, CA, USA) equipped with a diode array detector. Separation was obtained on an Agilent Poroshell 120 phenyl-hexyl column (2.1 mm x 250 mm, 2.7 μm) with a linear gradient consisting of water (A) and acetonitrile (B) both buffered with 20 mM formic acid, starting at 10 % B and increased to 100 % in 15 min where it was held for 2 min, returned to 10 % in 0.1 min and remaining for 3 min (0.35 mL/min, 60 °C). An injection volume of 1 μl was used. MS detection was performed in positive detection mode on an Agilent 6545 QTOF MS equipped with Agilent Dual Jet Stream electrospray ion source with a drying gas temperature of 250 °C, gas flow of 8 ml/min, sheath gas temperature of 300 °C and flow of 12 ml/min. Capillary voltage was set to 4000 V and nozzle voltage to 500 V. Mass spectra were recorded at 10, 20 and 40 eV as centroid data for m/z 85–1700 in MS mode and m/z 30–1700 in MS/MS mode, with an acquisition rate of 10 spectra/s. Lock mass solution in 70:30 methanol:water was infused in the second sprayer using an extra LC pump at a flow of 15 ml/min using a 1:100 splitter. The solution contained 1 M tributylamine (Sigma-Aldrich) and 10 M Hexakis (2,2,3,3-tetrafluoropropoxy)phosphazene (Apollo Scientific Ltd., Cheshire, UK) as lock masses. The $[\text{M}+\text{H}]^+$ ions (m/z 186.2216 and 922.0098 respectively) of both compounds was used.

2.7. Preparative isolation of acurin A

A. aculeatus was cultivated in 200 petri dishes with solid MM and incubated at 30 °C for five days. The plates were harvested and extracted twice overnight with ethyl acetate (EtOAc) with 1 % formic acid. The extract was filtered and concentrated in vacuo. The combined extract was dissolved in methanol (MeOH) and H₂O (9:1), and an equal amount of heptane was added where after the phases were separated. To the MeOH/H₂O phase H₂O was added to a ratio of 1:1, and metabolites were then extracted with dichlormethane (DCM). The phases were then concentrated separately in vacuo. The DCM phase was subjected to further purification on a semi-preparative HPLC, a Waters 600 Controller with a 996 photodiode array detector (Waters, Milford, MA, USA). This was achieved using a Luna II C18 column (250 mm x 10 mm, 5 μm , Phenomenex), a flowrate of 4 mL/min and an isocratic run at 40 % acetonitrile for

18 minutes. 50 ppm TFA was added to acetonitrile of HPLC grade and Milli-Q-water. This yielded 4.4 mg of acurin A.

2.8. NMR and structural elucidation

The 1D and 2D spectra were recorded on a Unity Inova-500 MHz spectrometer (Varian, Palo Alto, CA, USA). Spectra were acquired using standard pulse sequences and ^1H spectra as well as DQF-COSY, NOESY, HSQC and HMBC spectra were acquired. The deuterated solvent was DMSO- d_6 and signals were referenced by solvent signals for DMSO- d_6 at $\delta_{\text{H}} = 2.49$ ppm and $\delta_{\text{C}} = 39.5$ ppm. The NMR data was processed in Bruker Topspin 3.1. Chemical shifts are reported in ppm (δ) and scalar couplings in hertz (Hz). The sizes of the J coupling constants reported in Table S5 are the experimentally measured values from the spectra. There are minor variations in the measurements which may be explained by the uncertainty of J . The structural elucidation can be found in the supplementary material where NMR data for acurin A are listed in Table S5.

3. Results and Discussion

3.1. Acurin A – a novel PK-NRP hybrid compound from *Aspergillus aculeatus*

During our previous investigations of the 6-MSA synthase BGC, we discovered that *A. aculeatus* also produces a set of two minor compounds, both having the exact mass of 373.1889 Da (Petersen et al., 2015) corresponding to the elementary composition $\text{C}_{21}\text{H}_{27}\text{NO}_5$. Dereplication did not find a match from any known compounds with the same elementary composition present in Antibase, Reaxys, or our in-house collection of metabolites. The most significant (slightly earlier eluting) compound was isolated, and the structure was elucidated based on 2D NMR spectroscopy, see Table S5. This analysis revealed a compound likely to be a highly reduced polyketide coupled to an amino acid via a peptide bond, see Fig. 1A. Searching for the structure revealed that it had previously been reported in a Japanese patent (Nagata et al., 2012), however without establishment of any stereochemistry, which we report here. We have therefore named the compound acurin A. The PK-derived core carbon backbone part of acurin showed to be highly comparable to motifs previously observed in, e.g., lucilactaene produced by a *Fusarium* sp. (Kakeya et al., 2001), epolactaene from a marine *Penicillium* sp. (Kakeya et al., 1995), and most notably the mycotoxin fusarin C from various species of *Fusarium* (Cantalejo et al., 1999; Díaz-Sánchez et al., 2012; Niehaus et al., 2013; Song et al., 2004; Steyn and Vleggaar, 1985), see Fig. 1. Amongst these compounds, the

biosynthesis of fusarin C is best characterized. Hence, based on isotope labeling and genetic experiments, the scaffold of fusarin C appears to be made by the incorporation of a C4-unit amino acid, most likely L-homoserine, into the PK moiety of the fusarin C scaffold (Song et al., 2004). In agreement with this model, identification and analysis of the BGC responsible for fusarin C production demonstrated that it contained a gene, *fus1*, encoding a PKS-NRPS hybrid enzyme (Niehaus et al., 2013). The structural similarity to fusarin C suggests that acurin A may also originate from a PK and amino acid fusion. Assuming that the amino acid is L-serine (see later) the stereochemistry of the lactam ring can be established, since we observed an NOE correlation between the protons at positions C-15 and C-20. We note that this is similar to that observed for fusarin C (Fig. 1).

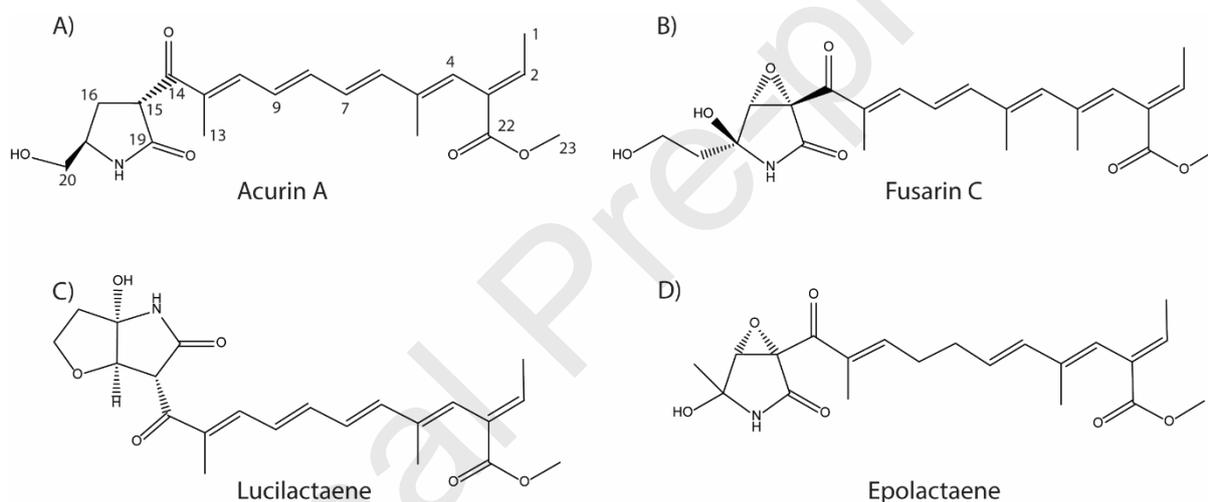


Fig. 1. Acurin A related compounds. A) Proposed structure of acurin A ($C_{21}H_{27}NO_5$, exact mass 373.1889 Da), assuming that L-serine is fused with the PK backbone. B) Structure of fusarin C ($C_{23}H_{29}NO_7$, exact mass 431.1944 Da). C) Structure of lucilactaene ($C_{22}H_{27}NO_6$, exact mass 401.1838 Da) D) Structure of epolactaene ($C_{21}H_{27}NO_6$, exact mass 389.4480 Da).

3.2. Identification of putative PKS and NRPS genes for the production of acurin

Based on the hypothesis that acurin is a fusion of a PK and an amino acid, we examined the genome sequence of *A. aculeatus* for the presence of a Fus1 homolog by a BlastP search to identify a BGC responsible for acurin production. Surprisingly, the protein with the highest degree of identity, 54 % (65 % query coverage), was a PKS (protein JGI ID 48) rather than a PKS-NRPS. This PKS was predicted to contain the following domains: ketosynthase (KS) - acyltransferase (AT) - dehydratase (DH) - methyltransferase (MT) - ketoreductase (KR) - acyl-carrier protein (ACP). The KS-AT-ACP are the basic domains for a non-reducing PKS, and

the presence of DH and KR domains, shows the PKS is a reducing type. The MT domain points to methylation of the scaffold. Strikingly, it did not appear to contain a release domain or an NRPS module (Marchler-Bauer et al., 2015). We, therefore, inspected the locus containing the gene (AACU_48) encoding this PKS for other genes that could be relevant for acurin formation. Initially, the analysis identified a putative NRPS gene (AACU_122295) with start-codon located approximately 31 kb upstream of the PKS encoding gene. The result indicates that the scaffold for acurin formation potentially could be synthesized by unlinked PKS and NRPS enzymes. The NRPS encoded by AACU_122295 showed only 37 % identity to the NRPS part of the Fus1 PKS-NRPS fusion protein, see Table 1 (Bushley and Turgeon, 2010; Lawrence et al., 2011). Importantly, the predicted NRPS contained a reductase domain for the potential release of a PKS-NRPS hybrid molecule.

Table 1. Genes surrounding *acrA* encoding the PKS involved in the biosynthesis of acurin, and the corresponding homologs in the *F. fujikuroi* fusarin C BGC

Protein/transcript AACU_ID (JGI)	Protein/transcript Aspac1_ID	Predicted Function	Gene name in <i>A. aculeatus</i> ^a	Homolog in <i>F. fujikuroi</i> fusarin C BGC ^b	Gene name in <i>F. fujikuroi</i>	Identity in %
1857757	-	Unknown	-	-	-	-
122297	62005	Endopeptidase	-	10055/11839	<i>fus4</i>	-
62004	-	CYP450	<i>acrF</i>	-	-	-
122295	773	NRPS	<i>acrB</i>	10058/11836	<i>fus1</i> *	37%
1857753	45009	Unknown	-	-	-	-
122294	31172	Dioxygenase	-	-	-	-
79887	31451	Transporter	-	-	-	-
122290	45006	TF	<i>acrR</i>	-	-	-
122285	53353	FAD Oxidoreductase	<i>acrE</i>	-	-	-
122283	61997	Methyl transferase	<i>acrG</i>	10050/11844	<i>fus9</i> *	46%
45003	61996	CYP450	<i>acrD</i>	10051/11843	<i>fus8</i> *	58%
		Aldehyde				
1903457	31526	Dehydrogenase	-	10052/11842	<i>fus7</i>	57%
1881869	-	MFS- transporter	-	10053/11841	<i>fus6</i>	47%
61993	53349	α/β hydrolase	<i>acrC</i>	10057/11837	<i>fus2</i> *	59%

48	-	PKS	<i>acrA</i>	10058/11836	<i>fus1</i> *	54%
61991	-	Unknown	-		-	-
1881866	-	β -transducin-like protein	-		-	-
1881865	-	Dehydrogenase	-		-	-

^aNames as defined and summarized in Figure.

^bTo denote the corresponding gene, the ID number is initiated with FFUJ_ (Niehaus et al., 2013) or Fufu. Different IDs for same protein are separated by / in this Table.

*Essential for fusarin C biosynthesis (Niehaus et al., 2013).

3.3. Separate PKS- and NRPS enzymes synthesize the acurin A core structure

To investigate whether AACU_48 (PKS) and AACU_122295 (NRPS) are involved in acurin production, the two genes were deleted. Conveniently, we previously used CRISPR/Cas9 to establish a *pyrG* marker in *A. aculeatus* (Nødvig et al., 2015), and we used this strain and CRISPR/Cas9 to construct an NHEJ deficient strain to increase the efficiency of gene-targeting experiments further. Specifically, we used an equivalent CRISPR/Cas9 vector with a protospacer targeting *akuA* accompanied by gene-targeting substrate employing a heterologous *pyrG* as selection marker, see Materials and Methods for details, Table S1 for genotypes, Table S3 and Fig. S1 & S2 for the strain validation. The resulting *akuA*Δ deletion strain, ACU10, did not produce a visible phenotype on solid minimal medium as compared to the wild-type strain. Next, in the *akuA*Δ *pyrG*1 strain background, we individually deleted AACU_48 and AACU_122295, see Materials and Methods and Figure S3 and S4, and analyzed the resulting mutant strains on solid minimal medium. Extracts from the resulting two strains were analyzed by ultra-high-performance liquid chromatography (UHPLC) coupled with diode array detection (DAD) and high-resolution mass spectrometry (HRMS) and compared with the corresponding chromatogram obtained with the reference strain. As expected, with the reference strain, we observed a significant but low-intensity peak representing acurin A. In contrast, with the AACU_48 and AACU_122295 deletion strains, the production of acurin A was abolished, see Fig. 2. Based on these findings, we, therefore, named the two genes *acrA* (AACU_48) and *acrB* (AACU_122295), respectively. Moreover, together with the bioinformatics analyses, these results strongly suggest that *acrA* encodes a PKS required for the biosynthesis of the PK moiety of acurin, and *acrB* encodes an NRPS, which catalyzes the fusion of an amino acid to the PK backbone of acurin and anticipated releases the primary precursor hybrid product. Further bioinformatics analyses showed that no epimerase domain is present in the NRPS. This supports our conclusion that L-serine is incorporated into not only acurin A, but also the slightly later elution isomer with the same elemental composition (Fig. 2). It therefore seems likely that the later eluting minor isomer of acurin A has the opposite stereochemistry at position-15.

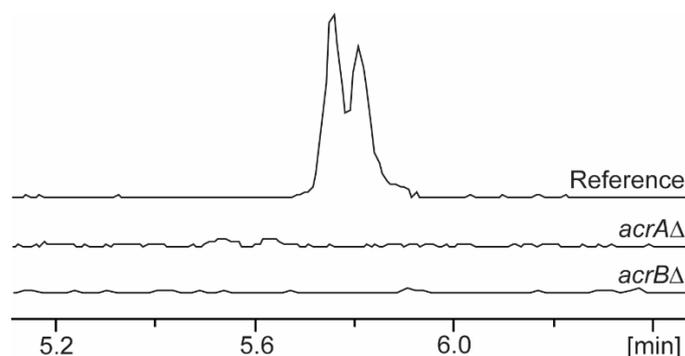


Fig 2. UHPLC-DAD-HRMS analysis of two *A. aculeatus* deletion strains and the reference strain cultivated on YES media. Traces represent extracted ion chromatograms (EICs) @ $m/z = 374.1962 \pm 0.002$ for the acurin A. Production of acurin A (and a likely slightly later eluting minor isomeric compound) was abolished when the genes encoding the PKS (*acrA*) and the NRPS (*acrB*) were deleted.

3.4. Uncoupled polyketide and non-ribosomal peptide synthetase activity is conserved in a fungal clade in section *Nigri*.

To our knowledge, fungal 3-acyl 2-pyrrolidone SM scaffolds described prior to this study are synthesized by a PKS-NRPS hybrid enzyme. Since the acurin A-like compounds are synthesized in genetically diverse fungi, we explored the possibility that BGCs for production of an identical or similar SM scaffold as the one used for acurin A production are present in other fungi. We therefore examined all available fungal genomes in the MycoCosm database, JGI, as well as NCBI using a BlastP search and the Acurin A producing PKS as a query.

Strikingly, this analysis revealed that nine other species within *Aspergillus* section *Nigri* had close homologs to AcrA as well as a separated AcrB (i.e. *A. fijensis*, *A. brunneoviolaceus*, *A. aculeatinus*, *A. violaceofuscus*, *A. japonicus*, *A. indologenus*, *A. uvarum*, *A. ellipticus* and *A. heteromorphus*), see Fig. S8. Moreover, these ten species are split in two clades in the overall species phylogeny within section *Nigri* (Vesth et al., 2018). Moreover, the species that are grouping in subclades share the organization of the majority of genes within the putative BGC spanning from the genes encoding *acrA* and *acrB*, whereas species in the most distant subclade, e.g. *A. uvarum*, showed an alternative organization in the BGC. This finding of multiple species having the BGC also rules out that the separation genes encoding the PKS and NRPS were an artefact from sequencing and genome assembly. Among the top hits from Section *Nigri*, we identified numerous PKS-NRPS hybrids, including Fus1 homologs. Similarly, when using Fus1 as a query the best hits were predominantly PKS-NRPS hybrids, except for the AcrA-like PKSs from *Aspergillus* section *Nigri*. This was also the outcome from an AcrB-based BlastP search.

Notably, coordinating catalytic moieties is favorable as it may support channeling effects to increase catalytic efficiency (Huang et al., 2001). However, the possibility exists that AcrA and AcrB achieve the same catalytic setup by forming a physical complex. This view is supported by the fact that the NRPS, but not the PKS, contains a recognized domain that provides a product-release mechanism. It was therefore of interest to us whether AcrA and AcrB enzymes represent an evolutionary stage that preceded the emergence of PKS-NRPS fusion enzymes or that fission occurred from the PKS-NRPS fusion after a horizontal gene transfer event to an ancestor of the *A. aculeatus* group.

In order to better understand the origin and evolution of the putative acurin A producing BGC, in particular for AcrA and AcrB, we conducted a comprehensive analysis of clustering and gene family evolution (Fig. S9) using a local database of 475 publicly available fungal genomes (Table S4). The functionally characterized fusarin C BGC and potential acurin A BGC were used as independent queries in a cluster discovery algorithm described previously (Reynolds et al., 2017). Seventeen BGCs of similar composition were recovered in Sordariomycetes, Dothideomycetes, Leotiomycetes, and Eurotiomycetes. Among recovered clusters, eight genes were widely conserved and shared by fusarin C and acurin A BGC loci. Several signature genes accompanied few or no other homologs of BGC genes in 5 Dothideomycetes and 3 Leotiomycetes and one Sordariomycete. Out of the 17 total BGCs, the acurin BGC is the only large cluster with separately encoded PKS and NRPS enzymes.

All query genes were subjected to maximum likelihood analyses, and Hybrid KS and AMP binding domain phylogenies were inferred separately. These analyses revealed that fusarin C and acurin BGCs form a clade along with similar gene clusters in a few distantly related fungi. A homologous cluster in *Metarrhizium* spp (Krasnoff et al., 2006), which are known to produce 7-desmethyl analogues of fusarin C, and in *Ophiocordyceps unilateralis* suggests that this may be the ancestral product of these clusters in Hypocreales. Most genes that are found in both acurin and fusarin C BGCs have a similar phylogenetic history among those genes that occur in the BGCs. This is consistent with the genes in the gene cluster evolving in a coordinated fashion. Despite this coordinated evolution, there is a minimal correspondence between the evolution of these genes and that of the species included in the database. Sporadic horizontal gene transfer might explain the sparse and distant distribution of homologous gene clusters, but there was no strong consistent signal for any specific transfer events. The restriction of the fusarin C homologs in the acurin BGC to *Aspergillus* section Nigri suggests a fusarin C-like

cluster was recently acquired in *Aspergillus*; however, the AMP binding domain and MFS transporter (AACU_1881869) phylogenies, in which *A. aculeatus* is sister to another Eurotiomycete species, are also consistent with a longer history in Eurotiomycetes. Acurin BGC genes that are not homologous to fusarin C BGC genes are more broadly found among *Aspergillus* spp.

The distribution and pattern of coordinated evolution among seven genes in these clusters suggests the ancestral state is a seven gene BGC, which broke down or rearranged in alternate lineages following its dispersal. Similarly, a parsimonious reconstruction suggests the ancestral state of the PKS/NRPS signature gene was a hybrid gene, which either fragmented (in *Aspergillus* section Nigri) or was degraded to a polyketide synthase in lineages (e.g., *Zymoseptoria*) where no homologous AMP binding domain was found. The AMP binding domain phylogeny roughly agrees with the Hybrid KS phylogeny, although the independent NRPS in *Aspergillus* section Nigri (e.g., Aspac1_802, Aspac1_773) do not agree with the *A. aculeatus* placement seen in the Hybrid KS phylogeny and most other gene phylogenies, see Figure 3. This conflict in gene phylogenies could be phylogenetic artifacts of accelerated evolution that followed the gene fragmentation and also a subsequent domain duplication, which might have relaxed selection on one or both of these paralogs.

Other notable gene-cluster evolution events include a rearrangement, possibly by a rolling circle transposition that gave rise to the fusarin C gene cluster in *Fusarium* with its one unique gene (*fus5*) at the recombination breakpoint, and a gene-cluster expansion in *Trichoderma*.

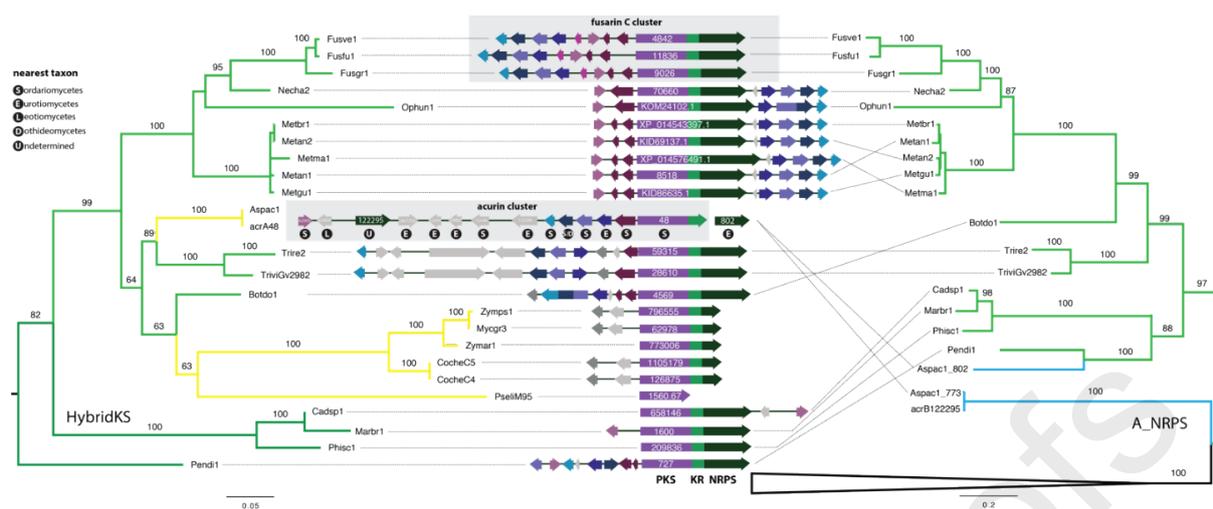


Fig. 3. Evolution of the acurin BGC. A maximum-likelihood phylogeny of the Hybrid KS domain (left) corresponds to the purple component of the PKS/NRPS signature gene in clusters (center). The maximum likelihood phylogeny (right) of the AMP binding domain corresponds to the dark green domain in the PKS/NRPS signature genes. Fusarin C and acurin gene clusters are shaded, and filled circles below the genes indicate the fungal class of the nearest sequence to acurin gene homologs. The clusters presented represent the total homologous clusters in a database of 475 publicly available fungal genomes (Table S4).

3.5 The acurin-associated BGC contains genes encoding tailoring enzymes and a transcription factor.

For the formation of acurin A, the primary PK-NRP product arising from *AcrA* and *AcrB* would likely need to be modified by a set of tailoring enzymes. We, therefore, made a careful inspection of the locus harboring *acrA* and *acrB* for other activities that could contribute to the biosynthesis. This analysis uncovered thirteen genes encoding potential SM biosynthesis activities, e.g., putative methyltransferases, oxidoreductases, cytochrome P450s. Thus, *acrA* and *acrB* would likely be part of a larger BGC.

We reasoned that the acurin A and fusarin C BGCs likely share several genes, and since the *Fus1* containing BGC in *F. fujikuroi* IMI 58289 had been characterized previously by Niehaus et al. (Niehaus et al., 2013), we decided to first compare the gene arrangement at the *fus* locus to that of the *acr* locus (Fig. 4). From the study in *Fusarium*, nine genes surrounding a central PKS-NRPS hybrid were co-regulated, but only four genes were found to be essential for fusarin C biosynthesis, see Table 1. From our syntenic analysis, at least six gene products shared homologs between the BGCs.

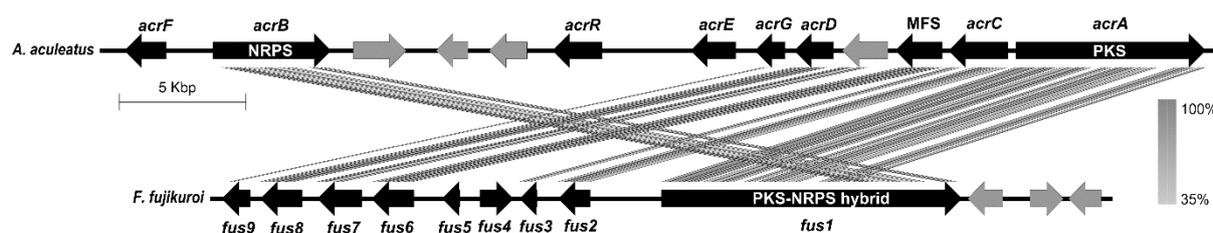


Fig. 4. Synteny between the fusarin C producing BGC and the BGC for producing acurin in *A. aculeatus*. The presumed acurin-producing BGC contains a gene encoding an iterative type I PKS and another gene encoding a single NRPS module. Neighboring the PKS- and NRPS-encoding genes are several genes representing various tailoring enzyme activities. Genes that we find are relevant for acurin production are shown by black arrows with their respective names above. Also, black arrows identify the proposed BGC for fusarin C, regardless of they being essential for production of fusarin C (Niehaus et al., 2013). Grey indicates an unrelated gene to the respective BGC (in *A. aculeatus* conformed by gene deletions (data not shown)).

Interestingly, the *acr* locus also contained a gene (AACU_122290) encoding a putative transcription factor (TF). To examine whether this TF had a regulatory impact on *acrA* and *acrB* and other genes in the locus, we eliminated AACU_122290, see Fig. S5. The mutation did not alter the morphology compared to the reference strain, but only by the abolition of acurin production, see Fig. S10. The result gave a direct link between the TF and acurin production, therefore, strongly indicating that the TF is involved in the activation of *acrA* and *acrB*. Hence, we named the gene AACU_122290, *acrR*.

Next, we performed RT-qPCR analysis to investigate the relative expression in the reference with the *acrRΔ* strain. Unfortunately, the relative expression levels of the genes in the *acr* locus between the reference and the *acrRΔ* strains were too low and unreliable for our measurement methods. Therefore, an overexpression (OE) construct for constitutive expression of the *acrR* was assembled to address whether the TF could activate the BGC and increase acurin production, see Materials and Methods section for details. As no defined integration site was established in *A. aculeatus* at this point, we decided to transform the *pyrG⁻* strain, which has a fully functional NHEJ pathway facilitating random integration of the *acrR* expression construct into the genome. The overexpression of *acrR* resulted in a significant phenotypic change with the production of a dark-red soluble pigment, and a decreased sporulation compared to the reference strain (Fig. S11). Four of the resulting transformants were analyzed by UHPLC-DAD-HRMS. All of the four mutants showed similar metabolite profiles, and interestingly, the metabolite profile revealed a drastic change in the metabolite profile when

compared to the reference strain (Fig. S12). The analysis also showed that the production of acurin was increased by the overexpression of *acrR* (Fig. S12).

We likewise performed an RT-qPCR to determine how the overexpression of *acrR* influenced the genes in the *acr* locus compared to gene-expression levels of a reference strain (Fig. 5 and S13). RNA was sampled after 3, 5, and 7 days growth on solid MM. We expected to see a significant change in the relative expression level in the *acr* locus between the two strains as the level of acurin significantly increased in the *acrR*-OE strain. Indeed, the expression levels of *acrA* and *acrB* were a 1000-fold increased at all measuring times in the *acrR*-OE strain. This trend was also observed for ten other genes in the locus. Four genes showed to be unaffected by the *acrR* overexpression, which indicated that they were not participating in the production of acurin.

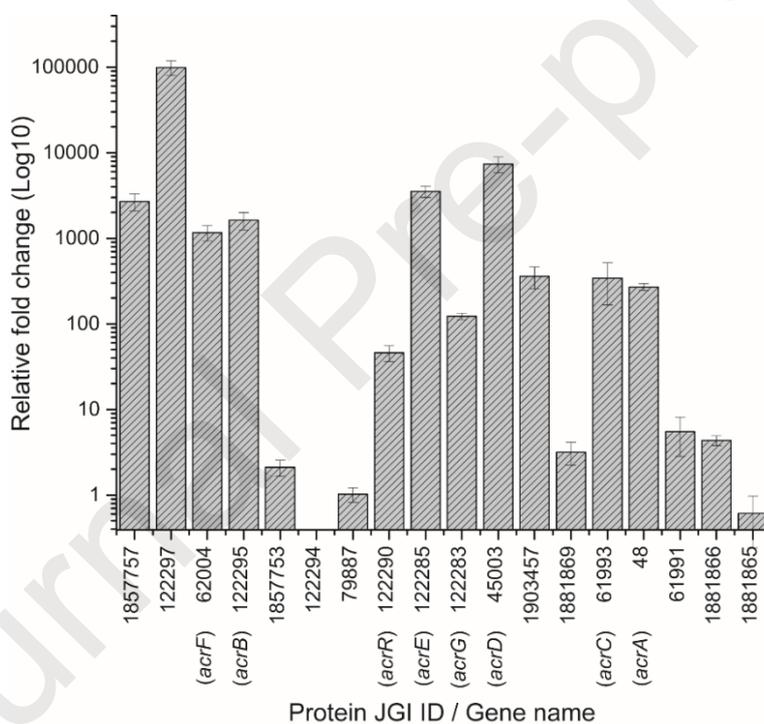


Fig. 5. Real-time quantitative PCR (RT-qPCR) data. Relative gene expression for the *acr* BGC between the reference strain and the *acrR*-OE strain measured at day 5, also see Fig. S13.

Based on our RT-qPCR analyses, it appears that AcrR regulates the BGC responsible for the formation of acurin and that the acurin responsible BGC may contain up to twelve genes. The massive increase in metabolites produced by the *acrR*-OE strain could be an artefact from

excess AcrR available in the nucleus, or it could indicate an impact on regulation of the secondary metabolism that is beyond the control of the acurin BGC for example as part of a broader and general response to an environmental cue (Fig. S12). The little influence on metabolism in the *acrR* Δ strain points to the first option, but it would be interesting to investigate in future studies.

3.6. Gene deletions identify genes in the *acr* biosynthetic gene cluster.

Deletion or overexpression of a TF-encoding gene regulating a BGC may influence the chromatin structure of the locus that harbors the BGC (Reyes-Dominguez et al., 2010; Woo Bok et al., 2009). Genes that are not part of the BGC, but happens to be situated close, or in the locus, may therefore also be transcribed differentially as the result of these genetic changes. To examine whether the genes that appear to be regulated by AcrR are directly involved in the production of acurin, we decided to delete the nine genes, which were affected by both the deletion and overexpression of *acrR*, see Fig. S6 and S7 for validation of strains.

Chromatograms from gene-deletion strains were compared to a reference for acurin A production levels and changes in other peaks. For differentiating peaks, the fragmentation pattern in mass spectra were analyzed in comparison to that of acurin A, since differences could originate from other compounds not relating to acurin A. Three of the deletion strains (i.e., AACU_122297 Δ , AACU_122294 Δ , AACU_1881869 Δ) were unaffected in acurin levels relative to the *acrR*-OE strain. We hypothesize that the induced expression of the three genes in the *acrR*-OE strain could be a result of the abnormal AcrR levels and its influence on the locus expression. Also, if AACU_1881869 encoding a putative MFS transporter has a role in the biosynthesis of acurin, either in extra- or intracellular transport, data would indicate that accumulation of acurin inside the fungus does not pose a threat to the fitness of the strain, or that this transport function was redundant.

The deletion of the five genes (AACU_61993/*acrC*, AACU_122285/*acrD*, AACU_62004/*acrE*, AACU_45003/*acrF*, and AACU_122283/*acrG*) resulted in the abolishment of acurin A production, as summarized in Figure 6 and Table 1, hence these genes were named according to the *acr* locus. To our surprise, four of these strains (*acrC* Δ , *acrD* Δ , *acrE* Δ , *acrF* Δ) revealed no accumulation of potential intermediates, which were not present in equivalent amounts in the reference strain. Since the amounts of acurin were much higher in the strain that overexpressed *acrR*, we

decided to also use this genetic feature to aid in the remaining gene-deletion strains. The ectopic integration of *acrA* was made in an NHEJ-proficient strain, however we needed a targeted approach based on the NHEJ-deficient strain. An equivalent plasmid was constructed where the *A. nidulans* integration sites were substituted to an integration site 1, *ISI*, for *A. aculeatus*. This locus *ISI* was placed between the genes AACU_117954 and AACU_59915, which encode putative a xylose transporter and RNA-binding protein/translational repressor, respectively. Therefore, spores from the gene-deletion strains *acrA-G* were cultivated on 5-FOA enriched medium, and for each gene deletion a uridine-requiring *pyrG* pop-out recombinant strain was selected as host strain for the subsequent integration of *acrR* in *ISI*. All transformants had the same extreme phenotype as the ectopic integration of *acrR*, and one representative colony from each strain was re-streaked, and then inoculated for extraction of metabolites and analysis by UHPLC-DAD-HRMS. Unfortunately, these strain did not show any accumulation of intermediates relating to acurin A. There are several plausible reasons for not detecting any compounds that appear in the deletion strains of *arcC*, *acrD*, *acrE*, and *acrF*. First, intermediate compounds could be unstable and, therefore, untraceable in the extracts of this genetic background displaying high complexity. Second, two or more of the enzymes in the pathway could interact physically during the reactions, which would mean that one missing enzyme would result in no formation of the enzyme complex.

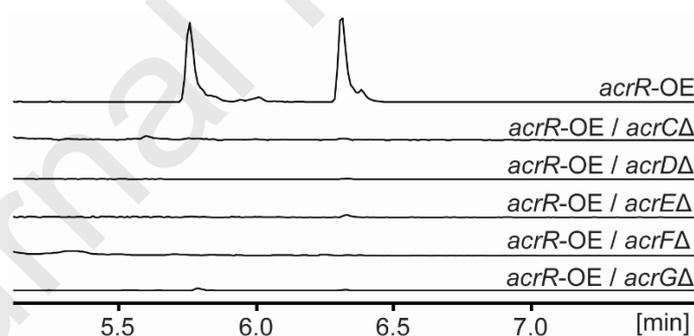


Fig. 6. UHPLC-DAD-HRMS analysis of *A. aculeatus* deletion strains cultivated on YES media. Traces represent EICs @ $m/z = 374.1962 \pm 0.005$ for acurin A. It appears that acurin A eluting around 5.8 min in the case of the *acrR*-OE strain is produced in much larger quantities than the co-eluting isomer detected in the reference strain (Fig. 2). Instead a later eluting other isomer is now detected in high quantities. We speculate that the appearance of this more apolar isomer of acurin A most likely represents an isomer with a changed stereochemistry in one of the double bonds in the PK part of the molecule, since both compounds are detected as having the same mass and thereby elemental composition.

One of the deletion mutant strains, AACU_122283/*acrG* (Fig. S14), did however show an accumulating product that was not found in the *acrA*Δ strain.

3.7. Proposed biosynthesis of acurin A

The acurin A producing BGC consists of at least eight genes; the PKS (*acrA*), NRPS (*acrB*), TF (*acrR*), and five genes that were encoding characteristic SM tailoring-enzyme activities; hydrolase (*acrC*), oxidoreductase (*acrD*), CYP450 (*acrE*), CYP450 (*acrF*), and O-methyltransferase (*acrG*). Interestingly, five of the gene products AcrA, AcrB, AcrC, AcrD, AcrG displayed homology (ranging from 37-59 % identity) to gene products essential for the fusarin C biosynthesis (Table 1). Therefore, it is plausible that these enzymes carry out equivalent functions in their respective biosynthetic pathways. On the other hand, since there are structural differences between acurin and fusarin C, the respective BGCs or genomes need to contain genes that are only present in that specific BGC. One example is the epoxide formation in fusarin C, where the responsible gene product remains elusive, and it may be performed by a gene located outside of the *fus* locus. This molecular feature is not present in acurin A, and has either been lost or never transferred between the respective species. In acurin A biosynthesis, the activities of the PKS (AcrA) and NRPS (AcrB) are collectively responsible for the synthesis of the acurin A core structure with a heptaketide backbone covalently fused to an amino-acid moiety, which we speculate could be L-serine, since this is the simplest possible scenario and L-serine is an abundant amino acid in the cell. After the formation of the PK-NRP hybrid product, we envision that it is detached from the enzyme by reductive release to set up the formation of the lactam ring by Aldol condensation followed by AcrC (hydrolase) catalyzed water loss to generate a double bond in the ring. Although AcrC is the closest homolog to Fus2 (59 % identity) from *F. fujikuroi*, which Niehaus and co-workers claim is responsible for lactam formation in fusarin C (Niehaus et al., 2013), we cannot, with our data prove that this is the equivalent function in *A. aculeatus*. We hypothesize that the double bond is reduced followed by three oxidations at C-22 to generate the carboxylic acid moiety. This hypothesis may be supported by the fact that AcrD-F encodes putative monooxygenases and two CYP450s (Fig. 7). However, the carboxylic acid formation is also occurring in fusarin C synthesis where there are no close homologs to AcrE and AcrF. If these enzymes are participating in this series of reactions, the genes responsible could have been recruited from another BGC. Both *acrE* and *acrF* reside in the left part of the BGC outside the region with high synteny to the fusarin C BGC. Alternatively, the homologs to AcrC and AcrD are Fus2 and Fus8, respectively, and both are essential for fusarin C biosynthesis. Thus, the steps carried

out by the homologs in the two BGCs could be highly similar. However, resolving their biosynthetic roles solely by predicted enzyme function is complicated by the fact that all of the four AcrC-F are not well-characterized types of enzymes.

An O-methyltransferase is needed for the final step in the biosynthesis of acurin A, and AcrG performs this step backed by MS/MS verification of the accumulating compound (Fig. S15 and S16) and the reported activity of the homolog Fus9 in fusarin C biosynthesis (Niehaus et al., 2013). We, therefore, conclude that AcrG is the O-methyltransferase acting as the final step in the biosynthesis of acurin A (Fig. 7).

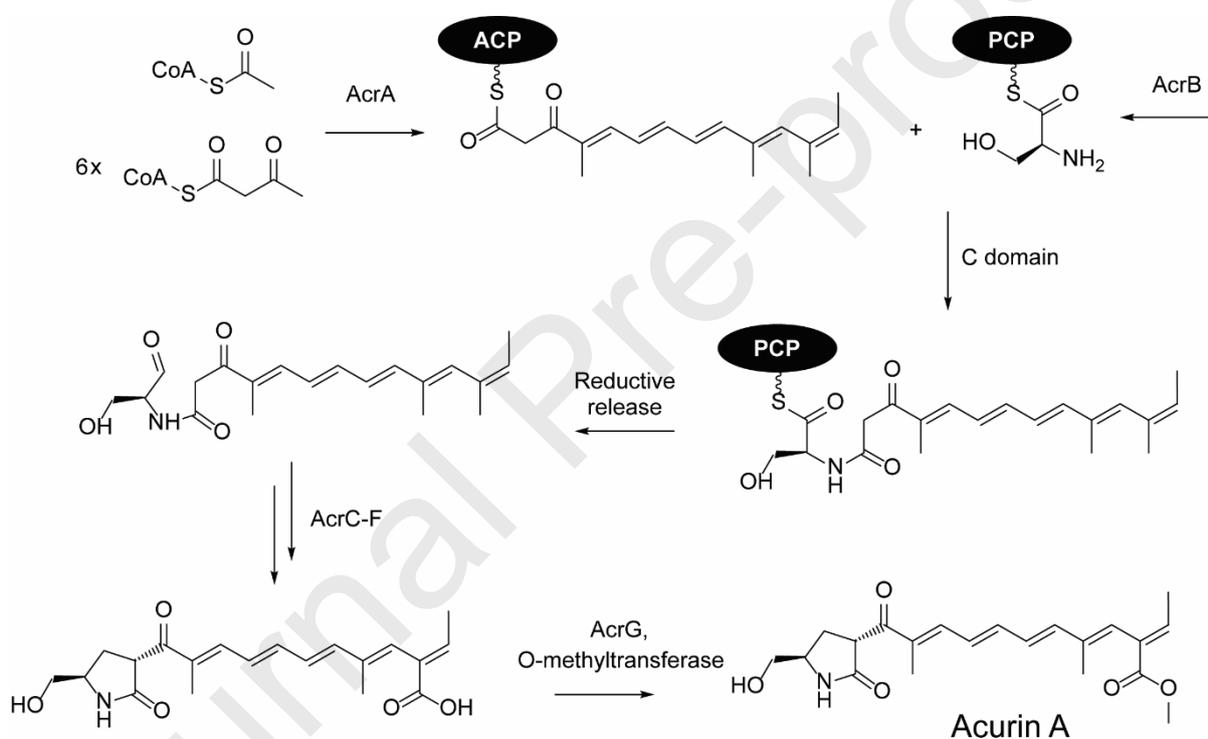


Fig. 7. The proposed biosynthetic pathway of acurin A.

Conclusions

In the present study, we demonstrate the power of employing analytical chemistry, comparative genomics, an efficient CRISPR/Cas9 system, and an NHEJ deficient strain to present a scheme of the route towards of acurin A in *Aspergillus aculeatus*. Acurin A shows high structural similarity to the mycotoxin fusarin C, differentiated by an epoxide functionality in fusarin C, another methylation pattern, and incorporation of a slightly different amino acid. Both BGCs seem to share the same putative enzymatic activities. Interestingly, we found that separate entities of PKS, AcrA, and NRPS, AcrB, fuse the heptaketide backbone to an L-serine moiety at the onset of the biosynthesis, while it has been shown that fusarin C is synthesized by a natural PKS-NRPS hybrid (Niehaus et al., 2013). The analyses from this study indicate that a fission of an ancestral gene encoding the PKS-NRPS hybrid resulted in separated *acrA* and *acrB* genes present in a few subclades in *Aspergillus* Section Nigri, and that AcrB involved in acurin A production may have been recruited for production of this compound at a later stage after the fission. Also, the activation of *acrA* and *acrB*, including the remainder of the BGC, were related to the transcription factor AcrR, which was encoded in BGC. At least eight genes (*acrA-G* and *acrR*) are vital for the biosynthesis of acurin A. The biological relevance and bioactivity of acurin A still needs to be investigated.

Acknowledgements

The work in the present study was supported by the Novo Nordisk Foundation, <http://www.novonordiskfonden.dk/en>, grants NNF12OC0000796 (MLN) and NNF15OC0016610 (PBW), and the National Science Foundation grant DEB- 1638999 (JCS), which we gratefully acknowledge.

References

- Andersen, R., Büchi, G., Kobbe, B., Demain, A.L., 1977. Secalonic Acids D and F Are Toxic Metabolites of *Aspergillus aculeatus*. UTC. <https://doi.org/10.1021/jo00422a042>
- Banerjee, D., Mondal, K.C., Pati, B.R., 2007. Tannase production by *Aspergillus aculeatus* DBF9 through solid-state fermentation. *Acta Microbiol. Immunol. Hung.* 54, 159–166. <https://doi.org/10.1556/AMicr.54.2007.2.6>
- Bennett, J.W., Klich, M., 2003. Mycotoxins. *Clin. Microbiol. Rev.* 16, 497–516. <https://doi.org/10.1128/CMR.16.3.497-516.2003>
- Blin, K., Shaw, S., Steinke, K., Villebro, R., Ziemert, N., Lee, S.Y., Medema, M.H., Weber, T., 2019. antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res.* 47, W81–W87. <https://doi.org/10.1093/nar/gkz310>

- Bushley, K.E., Turgeon, B.G., 2010. Phylogenomics reveals subfamilies of fungal nonribosomal peptide synthetases and their evolutionary relationships. *BMC Evol. Biol.* 10, 11. <https://doi.org/10.1186/1471-2148-10-26>
- Cantalejo, M.J., Torondel, P., Amate, L., Carrasco, J.M., Hernández, E., 1999. Detection of fusarin C and trichothecenes in *Fusarium* strains from Spain. *J. Basic Microbiol.* 39, 143–153. [https://doi.org/10.1002/\(SICI\)1521-4028\(199906\)39:3<143::AID-JOBM143>3.0.CO;2-U](https://doi.org/10.1002/(SICI)1521-4028(199906)39:3<143::AID-JOBM143>3.0.CO;2-U)
- Capella-Gutiérrez, S., Silla-Martínez, J.M., Gabaldón, T., 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25, 1972–1973. <https://doi.org/10.1093/bioinformatics/btp348>
- Clevenger, K.D., Bok, J.W., Ye, R., Miley, G.P., Verdán, M.H., Velk, T., Chen, C., Yang, K.H., Robey, M.T., Gao, P., Lamprecht, M., Thomas, P.M., Islam, M.N., Palmer, J.M., Wu, C.C., Keller, N.P., Kelleher, N.L., 2017. A scalable platform to identify fungal secondary metabolites and their gene clusters. *Nat. Chem. Biol.* <https://doi.org/10.1038/nchembio.2408>
- Díaz-Sánchez, V., Avalos, J., Limón, M.C., 2012. Identification and regulation of *fusA*, The polyketide synthase gene responsible for fusarin production in *Fusarium fujikuroi*. *Appl. Environ. Microbiol.* 78, 7258–7266. <https://doi.org/10.1128/AEM.01552-12>
- Edgar, R.C., 2010. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26, 2460–2461. <https://doi.org/10.1093/bioinformatics/btq461>
- Frisvad, J.C., Larsen, T.O., Thrane, U., Meijer, M., Varga, J., Samson, R.A., Nielsen, K.F., 2011. Fumonisin and ochratoxin production in industrial *Aspergillus niger* strains. *PLoS ONE* 6, 23496. <https://doi.org/10.1371/journal.pone.0023496>
- Gelderblom, W.C.A., Marasas, W.F.O., Steyn, P.S., Thiel, P.G., Van Der Merwe, K.J., Van Rooyen, P.H., Vlegaar, R., Wessels, P.L., 1984. Structure elucidation of fusarin C, a mutagen produced by *Fusarium moniliforme*. *J. Chem. Soc. Chem. Commun.* 122–124. <https://doi.org/10.1039/C39840000122>
- Hansen, B.G., Salomonsen, B., Nielsen, M.T., Nielsen, J.B., Hansen, N.B., Nielsen, K.F., Regueira, T.B., Nielsen, J., Patil, K.R., Mortensen, U.H., 2011. Versatile Enzyme Expression and Characterization System for *Aspergillus nidulans*, with the *Penicillium brevicompactum* Polyketide Synthase Gene from the Mycophenolic Acid Gene Cluster as a Test Case †. *Appl. Environ. Microbiol.* 77, 3044–3051. <https://doi.org/10.1128/AEM.01768-10>
- Hayashi, H., Furutsuka, K., Shiono, Y., 1999. Okaramines H and I, new okaramine congeners, from *Aspergillus aculeatus*. *J. Nat. Prod.* 62, 315–317. <https://doi.org/10.1021/np9802623>
- Holm, D.K., Petersen, L.M., Klitgaard, A., Knudsen, P.B., Jarczynska, Z.D., Nielsen, K.F., Gotfredsen, C.H., Larsen, T.O., Mortensen, U.H., 2014. Molecular and chemical characterization of the biosynthesis of the 6-MSA-derived meroterpenoid yanuthone

- D in *Aspergillus niger*. *Chem. Biol.* 21, 519–529.
<https://doi.org/10.1016/j.chembiol.2014.01.013>
- Huang, X., Holden, H.M., Raushel, F.M., 2001. Channeling of Substrates and Intermediates in Enzyme-Catalyzed Reactions. *Annu. Rev. Biochem.* 70, 149–180.
<https://doi.org/10.1146/annurev.biochem.70.1.149>
- Takeya, H., Kageyama, S.I., Nie, L., Onose, R., Okada, G., Beppu, T., Norbury, C.J., Osada, H., 2001. Lucilactaene, a new cell cycle inhibitor in p53-transfected cancer cells, produced by a *Fusarium* sp. [2]. *J. Antibiot. (Tokyo)* 54, 850–854.
<https://doi.org/10.7164/antibiotics.54.850>
- Takeya, H., Takahashi, I., Okada, G., Isono, K., Osada, H., 1995. Epolactaene, a Novel Neurotogenic Compound in Human Neuroblastoma Cells, Produced by a Marine Fungus. *J. Antibiot. (Tokyo)* 48, 733–735. <https://doi.org/10.7164/antibiotics.48.733>
- Katoh, K., Standley, D.M., 2013. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Mol. Biol. Evol.* 30, 772–780.
<https://doi.org/10.1093/molbev/mst010>
- Krasnoff, S.B., Sommers, C.H., Moon, Y.-S., Donzelli, B.G.G., Vandenberg, J.D., Churchill, A.C.L., Gibson, D.M., 2006. Production of Mutagenic Metabolites by *Metarhizium anisopliae*. *J. Agric. Food Chem.* 54, 7083–7088. <https://doi.org/10.1021/jf061405r>
- Lawrence, D.P., Kroken, S., Pryor, B.M., Arnold, A.E., 2011. Interkingdom gene transfer of a hybrid NPS/PKS from bacteria to filamentous ascomycota. *PLoS ONE* 6, e28231.
<https://doi.org/10.1371/journal.pone.0028231>
- Lee, C.F., Chen, L.X., Chiang, C.Y., Lai, C.Y., Lin, H.C., 2019. The Biosynthesis of Norsesquiterpene Aculenes Requires Three Cytochrome P450 Enzymes to Catalyze a Stepwise Demethylation Process. *Angew. Chem. - Int. Ed.* 18414–18418.
<https://doi.org/10.1002/anie.201910200>
- Marchler-Bauer, A., Derbyshire, M.K., Gonzales, N.R., Lu, S., Chitsaz, F., Geer, L.Y., Geer, R.C., He, J., Gwadz, M., Hurwitz, D.I., Lanczycki, C.J., Lu, F., Marchler, G.H., Song, J.S., Thanki, N., Wang, Z., Yamashita, R.A., Zhang, D., Zheng, C., Bryant, S.H., 2015. CDD: NCBI's conserved domain database. *Nucleic Acids Res.* 43, D222–D226.
<https://doi.org/10.1093/nar/gku1221>
- Nagata, H., Takeya, H., Konno, H., Kanazawa, S., 2012. JP2002322149-A ; JP4870276-B2: RKB-3384 analogs from *Aspergillus* species, and pharmaceuticals containing them.
- Nguyen, L.-T., Schmidt, H.A., von Haeseler, A., Minh, B.Q., 2015. IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. *Mol. Biol. Evol.* 32, 268–274. <https://doi.org/10.1093/molbev/msu300>
- Niehaus, E.M., Kleigrew, K., Wiemann, P., Studt, L., Sieber, C.M.K., Connolly, L.R., Freitag, M., Güldener, U., Tudzynski, B., Humpf, H.U., 2013. Genetic manipulation of the *fusarium fujikuroi* fusarin gene cluster yields insight into the complex regulation and

- fusarin biosynthetic pathway. *Chem. Biol.* 20, 1055–1066.
<https://doi.org/10.1016/j.chembiol.2013.07.004>
- Nielsen, M.T., Nielsen, J.B., Anyaogu, D.C., Holm, D.K., Nielsen, K.F., 2013. Heterologous Reconstitution of the Intact Geodin Gene Cluster in *Aspergillus nidulans* through a Simple and Versatile PCR Based Approach. *PLoS ONE* 8, 72871.
<https://doi.org/10.1371/journal.pone.0072871>
- Nørdvig, C.S., Nielsen, J.B., Kogle, M.E., Mortensen, U.H., 2015. A CRISPR-Cas9 system for genetic engineering of filamentous fungi. *PLoS ONE* 10, e0133085. <https://doi.org/10.1371/journal.pone.0133085>
- Nørholm, M.H.H., 2010. A mutant Pfu DNA polymerase designed for advanced uracil-excision DNA engineering. *BMC Biotechnol.* 10, 21. <https://doi.org/10.1186/1472-6750-10-21>
- Petersen, L.M., Hoeck, C., Frisvad, J.C., Gotfredsen, C.H., Larsen, T.O., 2014. Dereplication guided discovery of secondary metabolites of mixed biosynthetic origin from *Aspergillus aculeatus*. *Molecules* 19, 10898–10921.
<https://doi.org/10.3390/molecules190810898>
- Petersen, L.M., Holm, D.K., Gotfredsen, C.H., Mortensen, U.H., Larsen, T.O., 2015. Investigation of a 6-MSA Synthase Gene Cluster in *Aspergillus aculeatus* Reveals 6-MSA-derived Aculinic Acid, Aculins A-B and Epi-Aculin A. *ChemBioChem* 16, 2200–2204. <https://doi.org/10.1002/cbic.201500210>
- Price, M.N., Dehal, P.S., Arkin, A.P., 2010. FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. *PLoS ONE* 5.
<https://doi.org/10.1371/journal.pone.0009490>
- Puel, O., Galtier, P., Oswald, I.P., 2010. Biosynthesis and Toxicological Effects of Patulin. *Toxins* 2, 613–631. <https://doi.org/10.3390/toxins2040613>
- Reyes-Dominguez, Y., Woo Bok, J., Berger, H., Keats Shwab, ‡ E, Basheer, A., Gallmetzer, A., Scazzocchio, C., Keller, † Nancy, Strauss, J., 2010. Heterochromatic marks are associated with the repression of secondary metabolism clusters in *Aspergillus nidulans*. <https://doi.org/10.1111/j.1365-2958.2010.07051.x>
- Reynolds, H.T., Slot, J.C., Divon, H.H., Lysøe, E., Proctor, R.H., Brown, D.W., 2017. Differential Retention of Gene Functions in a Secondary Metabolite Cluster. *Mol. Biol. Evol.* 34, 2002–2015. <https://doi.org/10.1093/molbev/msx145>
- Robey, M.T., Ye, R., Bok, J.W., Clevenger, K.D., Islam, M.N., Chen, C., Gupta, R., Swyers, M., Wu, E., Gao, P., Thomas, P.M., Wu, C.C., Keller, N.P., Kelleher, N.L., 2018. Identification of the First Diketomorpholine Biosynthetic Pathway Using FAC-MS Technology. *ACS Chem. Biol.* 13, 1142–1147.
<https://doi.org/10.1021/acscchembio.8b00024>
- Samson, R. a., Houbraeken, J., Thrane, U., Frisvad, J.C., Andersen, B., 2010. Food and Indoor Fungi. *CBS Lab. Man. Ser.* 375–377.

- Secondary Metabolism Clusters - *Aspergillus aculeatus* ATCC16872 v1.1 [WWW Document], n.d. URL <https://mycocosm.jgi.doe.gov/pages/sm-clusters-summary.jsf?organism=Aspac1&genomes=Aspac1,Aspni5,Aspca3> (accessed 12.22.19).
- Smedsgaard, J., 1997. Micro-scale extraction procedure for standardization screening of fungal metabolite production in cultures. *J. Chromatogr. A* 760, 264–270. [https://doi.org/10.1016/S0021-9673\(96\)00803-5](https://doi.org/10.1016/S0021-9673(96)00803-5)
- Song, Z., Cox, R.J., Lazarus, C.M., Simpson, T.J., 2004. Fusarin C biosynthesis in *Fusarium moniliforme* and *Fusarium venenatum*. *ChemBioChem* 5, 1196–1203. <https://doi.org/10.1002/cbic.200400138>
- Steyn, P.S., Vleggaar, R., 1985. Biosynthetic studies on the fusarins, metabolites of *Fusarium moniliforme*. *J. Chem. Soc. Chem. Commun.* 1189–1191. <https://doi.org/10.1039/c39850001189>
- Suwannarangsee, S., Arnthong, J., Eurwilaichitr, L., Champreda, V., 2014. Production and characterization of multi-polysaccharide degrading enzymes from *Aspergillus aculeatus* BCC199 for saccharification of agricultural residues. *J. Microbiol. Biotechnol.* 24, 1427–1437. <https://doi.org/10.4014/jmb.1406.06050>
- Vesth, T.C., Nybo, J.L., Theobald, S., Frisvad, J.C., Larsen, T.O., Nielsen, K.F., Hoof, J.B., Brandl, J., Salamov, A., Riley, R., Gladden, J.M., Phatale, P., Nielsen, M.T., Lyhne, E.K., Kogle, M.E., Strasser, K., McDonnell, E., Barry, K., Clum, A., Chen, C., LaButti, K., Haridas, S., Nolan, M., Sandor, L., Kuo, A., Lipzen, A., Hainaut, M., Drula, E., Tsang, A., Magnuson, J.K., Henrissat, B., Wiebenga, A., Simmons, B.A., Mäkelä, M.R., de Vries, R.P., Grigoriev, I.V., Mortensen, U.H., Baker, S.E., Andersen, M.R., 2018. Investigation of inter- and intraspecies variation through genome sequencing of *Aspergillus* section *Nigri*. *Nat. Genet.* 50, 1688–1695. <https://doi.org/10.1038/s41588-018-0246-1>
- Woo Bok, J., Chiang, Y.-M., Szewczyk, E., Reyes-Dominguez, Y., Davidson, A.D., Sanchez, J.F., Lo, H.-C., Watanabe, K., Strauss, J., Oakley, B.R., Wang, C.C.C., Keller, N.P., 2009. Chromatin-level regulation of biosynthetic gene clusters. *Nat. Chem. Biol.* 5. <https://doi.org/10.1038/nchembio.177>

Author statement

Acurin A, a novel hybrid compound, biosynthesized by individually translated PKS- and NRPS-encoding genes in *Aspergillus aculeatus*

Peter B. Wolff^{a,1}, Maria L. Nielsen^{a,1}, Jason C. Slot^b, Lasse N. Andersen^a, Lene M. Petersen^a, Thomas Isbrandt^a, Dorte K. Holm^a, Uffe H. Mortensen^a, Christina S. Nødvig^a, Thomas O. Larsen^{*a}, Jakob B. Hoof^{*a}

^a Department of Biotechnology and Biomedicine (DTU Bioengineering), Technical University of Denmark, Søtofts Plads, Building 221/223, 2800 Kgs. Lyngby (Denmark). *E-mail: tol@bio.dtu.dk/jblni@dtu.dk.

^b Department of Plant Pathology, The Ohio State University, 481C Kottman Hall, 2021 Coffey Road, Columbus, OH 43210 (USA).

¹ These authors have contributed equally to this work.

Peter B. Wolff: Conceptualization, Writing - original draft - review & editing, Investigation, Methodology.

Maria L. Nielsen: Conceptualization, Writing - original draft, Investigation molecular biology, Methodology molecular biology.

Jason C. Slot: Writing - original draft, review & editing, Investigation bioinformatics, Methodology bioinformatics.

Lasse N. Andersen: Investigation molecular biology, Methodology molecular biology.

Lene M. Petersen: Investigation chemistry, Methodology chemistry.

Thomas Isbrandt: Investigation chemistry, Writing - review & editing.

Dorte K. Holm: Methodology molecular biology.

Uffe H. Mortensen: Writing - review & editing, Supervision molecular biology.

Christina S. Nødvig: Methodology molecular biology.

Thomas O. Larsen: Writing - original draft - review & editing, Funding acquisition, Supervision chemistry, Investigation, Methodology.

Jakob B. Hoof: Conceptualization, Writing - original draft - review & editing, Supervision molecular biology, Investigation, Methodology.

Acurin A, a novel hybrid compound, biosynthesized by individually translated PKS- and NRPS-encoding genes in *Aspergillus aculeatus*

Peter B. Wolff^{a,1}, Maria L. Nielsen^{a,1}, Jason C. Slot^b, Lasse N. Andersen^a, Lene M. Petersen^a, Thomas Isbrandt^a, Dorte K. Holm^a, Uffe H. Mortensen^a, Christina S. Nødvig^a, Thomas O. Larsen^{*a}, Jakob B. Hoof^{*a}

^a Department of Biotechnology and Biomedicine (DTU Bioengineering), Technical University of Denmark, Søtofts Plads, Building 221/223, 2800 Kgs. Lyngby (Denmark). *E-mail: tol@bio.dtu.dk/jblni@dtu.dk.

^b Department of Plant Pathology, The Ohio State University, 481C Kottman Hall, 2021 Coffey Road, Columbus, OH 43210 (USA).

¹ These authors have contributed equally to this work.

Highlights

- A biosynthetic gene cluster for the hybrid polyketide-nonribosomal peptide compound acurin A in *Aspergillus aculeatus*.
- The PKS and NRPS responsible for acurin A production are two individually transcribed entities.
- Phylogenetics suggest that the separated PKS and NRPS is a result from the fission of an ancestral PKS-NRPS hybrid.
- The biosynthetic capacity of *A. aculeatus* is facilitated by the use of CRISPR/Cas9 gene editing.