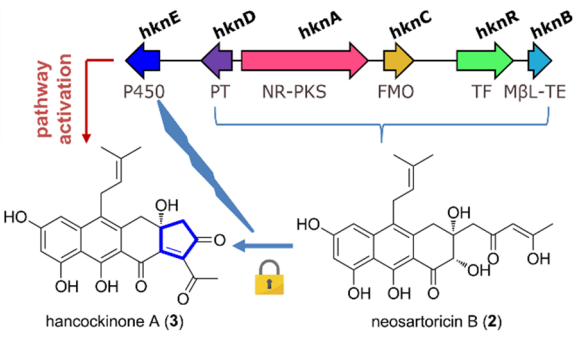
Genome Mining of *Aspergillus hancockii* Unearths Cryptic Polyketide Hancockinone A Featuring Prenylated 6/6/6/5 Carbocyclic Skeleton

Hang Li,[*,†,Δ] Si Shu,[†] John A. Kalaitzis,[‡,§] Zhuo Shang,[†] Daniel Vuong,[§] Andrew Crombie,[§] Ernest Lacey,[‡,§] Andrew M. Piggott,[*,‡] Yit-Heng Chooi[*,†]

[†]School of Molecular Sciences, The University of Western Australia, Perth, WA 6009, Australia

[‡]Department of Molecular Sciences, Macquarie University, Sydney, NSW 2109, Australia

[§]Microbial Screening Technologies Pty. Ltd., Smithfield, NSW 2164, Australia



ABSTRACT: Activation of a cryptic polyketide synthase gene cluster *hkn* from *Aspergillus hancockii* via overexpression of the gene cluster-specific transcription factor HknR led to the discovery of a novel polycyclic metabolite, which we named hancockinone A. The compound features an unprecedented prenylated 6/6/6/5 tetracarbocyclic skeleton and shows moderate antibacterial activity. Heterologous expression, substrate feeding and *in vitro* assays confirmed the role of cytochrome P450 HknE in constructing the five-membered ring in hancockinone A from the precursor neosartoricin B.

Polycyclic aromatic polyketides (PcAPs) are an important group of natural products that are of significant interest in drug discovery and include the well-known anticancer drug doxorubicin[1] and antibiotics of the tetracycline family[2] (Figure 1), both derived from actinobacteria. Fungi are also prolific producers of PcAPs, which have been shown to possess a diverse range of biological activities, including the antifungal griseofulvin,[3] the mycotoxin aflatoxin,[4] the meroterpenoids hypomycetin[5] and viridicatumtoxin,[6] as well as the antibacterial and anticancer viriditoxin[7] (Figure 1).

In fungi, PcAPs are typically synthesized by iterative non-reducing polyketide synthases (NR-PKSs) and further tailored by post-modification enzymes, such as cytochrome P450s, flavin-dependent monooxygenases (FMOs), methyltransferases, etc.[8] Fungal genome sequencing over the last decade has identified a large number of NR-PKS gene clusters potentially encoding the production of PcAPs,[9] but many are "silent" under laboratory culture conditions. This suggests that the biosynthetic potential of fungi for making diverse PcAPs is much higher than was initially appreciated and that targeted genome mining can be a fruitful approach for the discovery of bioactive PcAPs.

Neosartoricin (**1**), which exhibits immunosuppressive properties, is one such PcAP discovered from *Neosartorya fischeri* via a genome mining approach.[10] The previous study has demonstrated that the desacetyl precursor, neosartoricin B (**2**), is the direct metabolite biosynthesized by a gene cluster *nsc*, which encodes a NR-PKS, a FMO, a prenyltransferase and a metallo-β-lactamase-like thioesterase (MβL-TE) (Figure 2A). A follow-up study demonstrated that the core set of biosynthetic genes for **2** is conserved in several fungi that are pathogenic to humans and other animals.[11] Furthermore, some of those orthologous gene clusters encode additional post-modification enzymes, such as dioxygenases and alcohol dehydrogenases (e.g. in *Trichophyton rubrum*) (Figure 2A).[11] This suggests that homologous gene clusters could encode the production of derivatives of **2** with more complex structures. Nevertheless, the final metabolites encoded by these gene clusters have remained elusive.

During our recent in-depth genomic investigation of the biosynthetically talented Australian fungus *Aspergillus hancockii*,[12] we identified a putative NR-PKS gene cluster *hkn*. The *hkn* gene cluster shares high similarity with the *nsc* cluster responsible for the biosynthesis of **2**,[10-11] with the only difference being the presence of an additional cytochrome P450 gene *hknE* (Figure 2A and Table S1). The

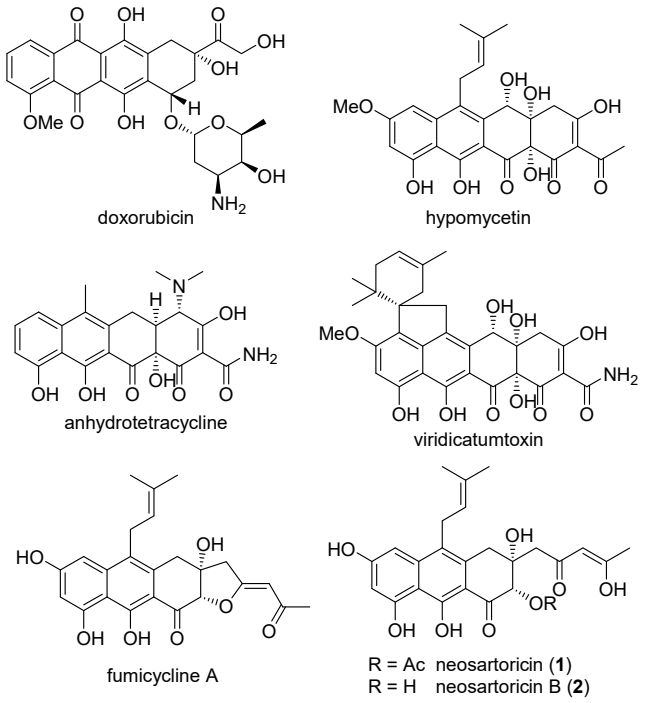


Figure 1. Representative tetracyclic aromatic polyketides.

close homology between the *hkn* and *nsc* clusters suggests that the *hkn* cluster may encode the production of **2**-like compounds. Interestingly, the production of such **2**-like metabolites by *A. hancockii* has not been observed in any of the various culture conditions tested.[12a]

To establish a link between the *hkn* cluster and its encoded metabolites, we overexpressed a putative Zn_2Cys_6 transcriptional regulator encoded by *hknR* in the cluster using the constitutive *A. nidulans gpdA* promoter. This gene cluster activation method has been successfully demonstrated in many fungi,[13] including the neosartoricin producer *N. fischeri*.[10] The resulting glufosinate-resistant transformants displayed yellow pigmentation on the reverse side of the agar plates (Figure S1). HPLC-DAD-MS analysis of the acetone extracts from the mutant grown in a stationary minimal medium liquid culture revealed the accumulation of a new peak **3** (*m/z* 423 $[M + H]^+$) with a highly characteristic UV-vis spectrum, which was absent in the wild-type (WT) strain (Figure 2B).

To isolate **3**, the *hknR* overexpression mutant was grown on jasmine rice (2.5 kg) for 3 weeks, as described previously for *A. hancockii*.[12] Subsequent extraction and purification afforded **3** (97 mg) as a red solid (Supporting Information). HRESI(+)MS analysis of **3** ($[M + H]^+$ *m/z* 423.1447) indicated a molecular formula $C_{24}H_{22}O_7$ containing four hydrogens and one oxygen fewer than **2**, and requiring 14 double bond equivalents (DBE), compared to 12 DBE in **2**. The UV-vis spectrum of an intensely purple-colored solution of **3** in MeCN revealed a large bathochromic shift from λ_{max} 405 nm in **2** to 511 nm in **3** (Figure 2C), suggesting the presence of significantly more extended conjugation. The ^1H (Figure S8) and ^{13}C (Figure S9) NMR spectra of **3** in DMSO-d_6 were similar to those of **2**,[11] with the most significant differences being the presence of only one set of NMR signals (i.e. no keto-enol tautomers), the absence of signals associated with oxymethine C-2 (δ_H 4.49/4.54; δ_C

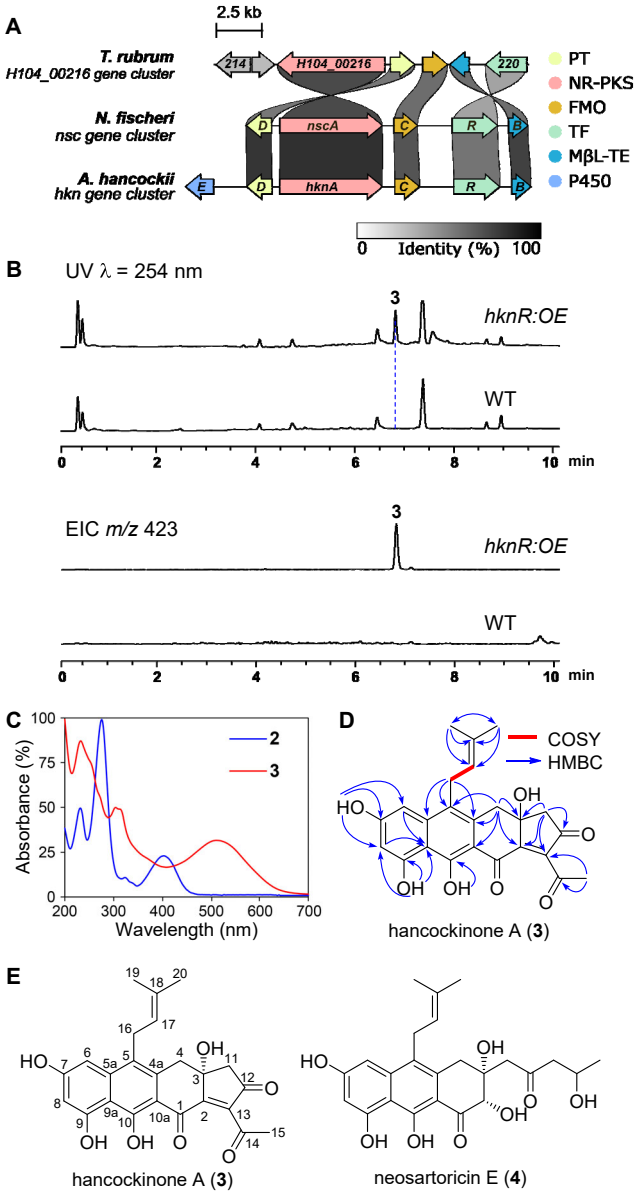


Figure 2. Overexpression (OE) of *hknR* in the *hkn* gene cluster led to the production of hancockinone A (**3**). (**A**) The *hkn* gene cluster compared to *nsc* cluster (homologues are in same colors). (**B**) LC-MS metabolite profile analysis of hknR-OE strains compared to WT. EIC, extracted ion chromatogram. (**C**) Comparison of the UV-vis spectra of **2** and **3**. (**D**) Selected 2D NMR correlations for **3**. (**E**) Structures of **3** and **4**.

76.4/76.2), hydroxy group 2-OH (δ_H 5.66/5.71) and enol/diketo olefin/methylene C-13 (δ_H 5.82/3.79; δ_C 102.6/58.7), and the presence of two additional olefinic carbon signals (δ_C 158.1 and δ_C 142.9) (Table S4). To account for the second additional DBE, we hypothesized that cyclization had occurred between C-2 and C-13 to form a five-membered ring, with concomitant dehydration yielding a tetrasubstituted double bond ($\Delta^{2,13}$). This hypothesis was supported by the presence of HMBC correlations from H-4α and H_2-11 to C-2 (δ_C 158.1) and from H-11β and H_3-15 to C-13 (δ_C 142.9) (Figure 2D, Table S5 and Figure S12). Thus, the planar structure of **3** was elucidated as shown in Figure 2E. The absolute configuration of C-3

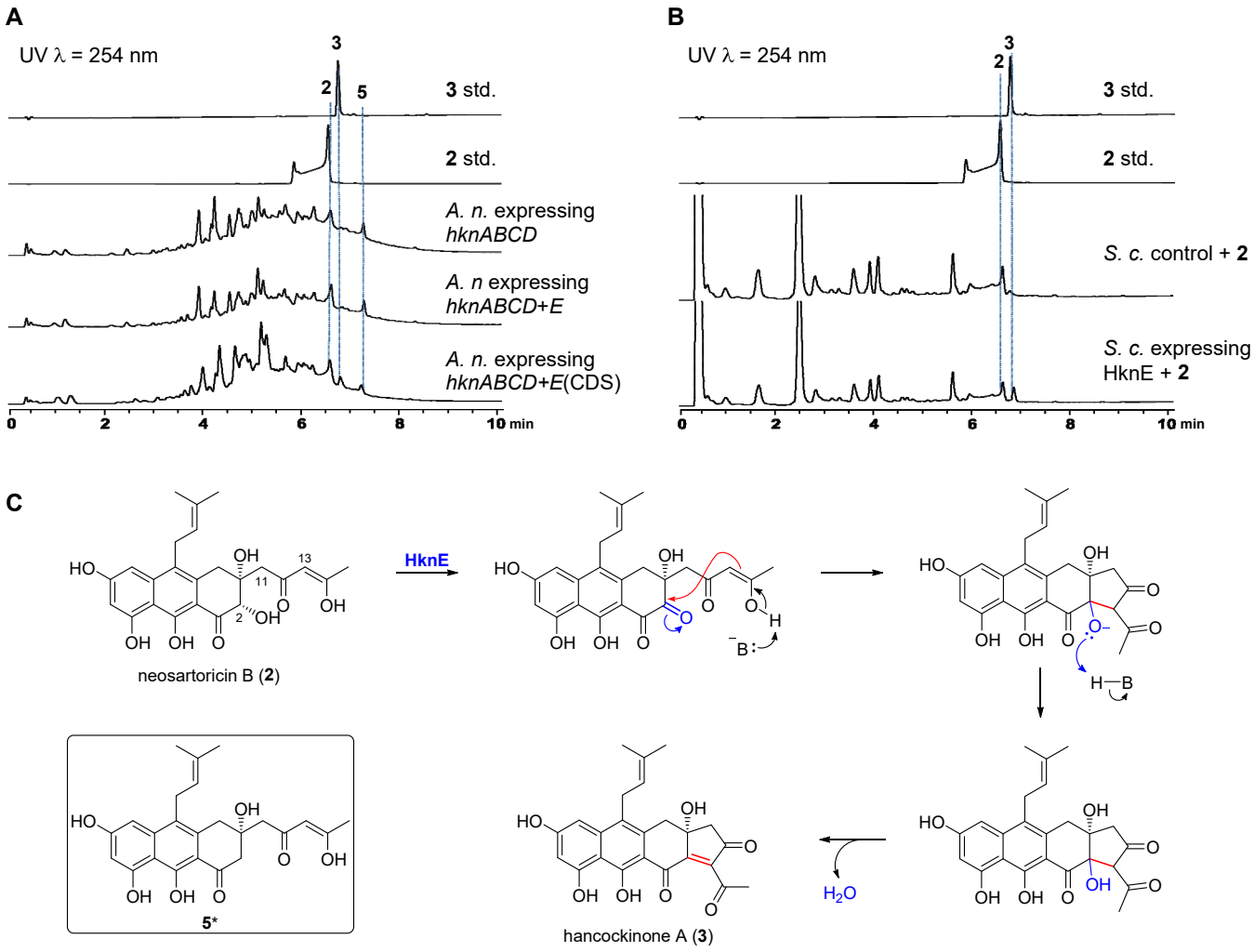


Figure 3. (**A**) HPLC traces demonstrating heterologous production of **3** in *A. nidulans* mutants (**B**) HPLC traces demonstrating *in vivo* biotransformation of **2** to **3** by *S. cerevisiae* BJ5464 expressing HknE. *S. c.* control means *S. cerevisiae* harboring pXP741 empty plasmid. The peak broadening observed in **2** standard is due to interconversion of keto-enol tautomers. (**C**) Proposed mechanism for the HknE-catalyzed cyclization of **2** to **3**. * Proposed structure of **5** (boxed) based on UV-vis and MS data.

was tentatively assigned as *S* based on its biosynthetic origin from **2**, as discussed below. Compound **3** structurally resembles fumicycline A (Figure 2), which was identified via activation of the *fcc* gene cluster (the orthologous cluster of *nsc*) in *A. fumigatus*.[14] Fumicycline A also harbors a 6/6/6/5 tetracyclic system, although the D ring in fumicycline A is a tetrahydrofuran, possibly formed from **2** via non-enzymatic cyclodehydration (Figure S4).[14] Notably, polyketides with a tetracarbocyclic 6/6/6/5 skeleton have not been previously reported. In addition to **3**, from the large-scale fermentation, we also isolated **1** and **2** along with a minor compound **4** (Figure S7), which we elucidated as the 13,14-dihydro derivative of **2**, herein named as neosartoricin E (Figure 2E, Figures S14-S19 and Table S5). The production of **4** is likely due to an endogenous ketoreductase encoded in *A. hancockii*. A similar metabolite with the ketone reduced to the alcohol has also been observed previously.[10]

To further validate that the putative *hkn* cluster is both required and sufficient for the biosynthesis of **3**, we employed a heterologous pathway expression approach. Considering the extensive overlap between the *hkn* and *nsc* clusters, we first expressed *hknABCD* (akin to *nscABCD* encoding the production of **2**) in a characterized host *A. nidulans* LO8030[15] under alcohol-inducible promoters using the previously established pYFAC fungal episomal vectors.[16] Expectedly, the production of **2** was detected in the extracts of *A. nidulans* expressing *hknABCD* (Figure S2), which corresponds to coexpression of *nscABCD* in *A. nidulans*.[11] In addition to **2**, a putative pathway intermediate **5** was also detected (Figure 3A). Based on MS analysis ([M + H]$^+$ *m/z* 427) and an almost identical UV-vis spectrum to **2** (Figure S2), we tentatively assigned the structure of **5** as 2-deoxyneosartoricin B (Figure 3C). This compound has been observed previously in *N. fischeri* overexpressing *nscR* and was proposed as the pathway intermediate before the FMO HknC hydroxylates **5** to **2**.[10]

Next, we coexpressed the P450 gene *hknE* together with *hknABCD*. While the corresponding transformant still produced **2** and **5**, compound **3** was not detected (Figure 3A). We proposed that this could be due to the introns in *hknE* failing to be spliced correctly in *A. nidulans* (Table S7),

which has been previously observed in other BGC heterologous expression studies.[17] To verify this hypothesis, we introduced intron-free *hknE* cloned from the cDNA synthesized from mRNA extracted from the *A. hancockii hknR* overexpression mutant, into *A. nidulans hknABCD.* Pleasingly, HPLC analysis confirmed the production of **3** by this mutant (Figure 3A).

To further confirm the function of the enzyme HknE in the biosynthesis of **3**, we expressed HknE in *Saccharomyces cerevisiae* JHY702.[18] The intron-free *hknE* was cloned under the ADH2 promoter into the pXP741 vector.[19] Upon feeding of **2** into a culture of this yeast strain expressing HknE, the consumption of **2** and the accumulation of **3** were observed after 48 h of incubation (Figure 3B). Microsome extracts of HknE were also prepared from *S. cerevisiae* expressing *hknE* for *in vitro* assay. Following incubation of **2** and NADPH with the microsomal HknE, the formation of **3** was clearly observed (Figure S3). These *in vivo* and *in vitro* results support the assertion that i) the *hkn* gene cluster is sufficient to biosynthesize **3**, and ii) HknE is the single enzyme required to convert **2** into **3**.

To explore possible mechanisms of HkmE-mediated five-membered ring formation, we tested the stability of **2** in a variety of common solvents, acids and bases. Interestingly, we found that **2** underwent almost complete conversion to **3** in both 10% triethylamine in acetonitrile and in 100% DMSO after stirring for 18 h at room temperature (Figures S5-S6). No transformation was observed in either aqueous 0.1% NaOH or 0.1% HCl. Based on these observations, it is likely that the conversion of **2** to **3** can be catalyzed by a weak base and is promoted in an oxidative environment, with the extended conjugation of **3** being a significant thermodynamic driving force for this reaction. Herein, we propose that HknE firstly oxidizes the hydroxy group at C-2 to a ketone intermediate. Then, an intramolecular aldol condensation forms a carbon-carbon bond between C-2 and C-13, possibly mediated by a basic residue in HknE, leading to the construction of the five-membered ring (Figure 3C). Subsequently, spontaneous dehydration leads to the conjugated cyclopentenone in **3**. To probe this plausible mechanism, we tested **4** as a potential substrate of HknE to explore whether we could capture a corresponding keto intermediate. However, when **4** was fed to yeast expressing *hknE*, a corresponding keto compound was not detected (Figure S25). This could be due to HknE exhibiting a high degree of selectivity for **2**. P450s have been widely shown to oxidize a secondary alcohol to the corresponding ketone, which then triggers or facilitates subsequent cyclization,[20] and HknE provides another example of P450-mediated strained ring construction via oxidation of a secondary alcohol. These results, taken together with the results of earlier studies,[8a, 10] have allowed us to propose a plausible biosynthetic pathway to **3** (Figure S4).

Considering PcAPs have been reported to have a wide range of bioactivities, we next screened **1**–**4** for *in vitro* antibacterial, antifungal, antiparasitic, cytotoxic and herbicidal activities. Compounds **1**–**3** exhibited moderate *in vitro* antibacterial activities against *Bacillus subtilis* (25, 12.5 and 25 μg/mL respectively) and *Staphylococcus aureus* (25, 6.3 and 25 μg/mL respectively), but no significant antifungal, antitumor or herbicidal activities were detected for any of the compounds (Table S6).

In conclusion, a cryptic NR-PKS gene cluster (*hkn*) from *A. hancockii* was demonstrated, via pathway activation in the native host, to encode the biosynthesis of a prenylated PcAP hancockinone A (**3**). The *hkn* cluster harbors an additional cytochrome P450-encoding gene *hknE* compared to the well-studied *nsc* cluster encoding the production of neosartoricin B (**2**). The cytochrome P450 HknE has been confirmed to catalyze an oxidative cyclization to convert the tricyclic **2** into an unprecedented 6/6/6/5 tetracarbocycle **3**. Using cblaster,[21] we identified a core set of biosynthetic genes, comprising NR-PKS, PT, FMO and MβL-TE, that is conserved in over two dozen fungi in the NCBI database (Figure S26), including those previously identified.[10-11] Notably, *hknE* is unique to *A. hancockii*. Several of these gene clusters encode additional post-modification enzymes, including methyltransferases, oxidoreductases and PLP-dependent enzymes (Figure S27). This suggests that more diverse prenylated PcAPs encoded in these fungi are yet to be characterized. Previous studies on such cryptic pathways via a heterologous expression method failed to identify the final encoded metabolites,[11] possibly because some of biosynthetic genes were not expressed in the heterologous host, as was initially the case in this study, and thus warrant reinvestigation. Our study also highlights the advantages of characterizing cryptic pathways via pathway activation in the native strain if the fungus is genetically tractable. Alternatively, if employing a heterologous biosynthesis approach, potential problems with intron-splicing should be considered if the product obtained was not as expected based on gene cluster analysis.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website.

Experimental details, spectroscopic data, supplementary figures and tables (PDF).

AUTHOR INFORMATION

Corresponding Authors

*H.L.: E-mail: hang.li@uwa.edu.au
*A.M.P.: E-mail: andrew.piggott@mq.edu.au
*Y.H.C.: E-mail: yitheng.chooi@uwa.edu.au

Present Addresses

[Δ] School of Pharmaceutical Sciences, Sun Yat-sen University, Guangzhou, Guangdong 510006, China; E-mail: lihang33@mail.sysu.edu.cn

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENT

We thank Samuele Sala (UWA) for acquisition of HR-ESI-MS data, Kymren Boelling-McDougall (MQ) for acquisition of UV-vis data and Mahmud Morshed (MQ) for acquisition of ECD data. This research was funded, in part, by the Australian Research Council (FT160100233 and DP210102180) and a Cooperative Research Centres Project (CRCPFIVE000119).

REFERENCES

1. a) Arcamone, F.; Cassinelli, G.; Fantini, G.; Grein, A.; Orezzi, P.; Pol, C.; Spalla, C., *Biotechnol. Bioeng.* **1969,** *11*, 1101-1110; b) Hutchinson, C. R., *Chem. Rev.* **1997,** *97*, 2525-2536.

2. Grossman, T. H., *Cold Spring Harb. Perspect. Med.* **2016,** *6*, a025387-a025387.

3. Cacho, R. A.; Chooi, Y.-H.; Zhou, H.; Tang, Y., *ACS Chem. Biol.* **2013,** *8*, 2322-2330.

4. Richard, J. L., *Toxin Rev.* **2008,** *27*, 171-201.

5. Breinholt, J.; Jensen, G.; Kjær, A.; Olsen, C.; Rosendahl, C.; Romerosa, A.; Robinson, W.; Roos, B.; Vallance, C.; Wood, B., *Acta Chem. Scand.* **1997,** *51*, 855-860.

6. a) Zheng, C.-J.; Yu, H.-E.; Kim, E.-H.; Kim, W.-G., *J. Antibiot.* **2008,** *61*, 633-637; b) Li, W.; Li, L.; Zhang, C.; Cai, Y.; Gao, Q.; Wang, F.; Cao, Y.; Lin, J.; Li, J.; Shang, Z.; Lin, W., *ACS Infect. Dis.* **2020,** *6*, 1759-1769.

7. a) Wang, J.; Galgoci, A.; Kodali, S.; Herath, K. B.; Jayasuriya, H.; Dorso, K.; Vicente, F.; González, A.; Cully, D.; Bramhill, D.; Singh, S., *J. Biol. Chem.* **2003,** *278*, 44424-44428; b) Kundu, S.; Kim, T. H.; Yoon, J. H.; Shin, H.-S.; Lee, J.; Jung, J. H.; Kim, H. S., *Int. J. Oncol.* **2014,** *45*, 2331-2340.

8. a) Li, Y.; Chooi, Y.-H.; Sheng, Y.; Valentine, J. S.; Tang, Y., *J. Am. Chem. Soc.* **2011,** *133*, 15773-15785; b) Chooi, Y.-H.; Tang, Y., *J. Org. Chem.* **2012,** *77*, 9933-9953.

9. a) Kjærbølling, I.; Vesth, T. C.; Frisvad, J. C.; Nybo, J. L.; Theobald, S.; Kuo, A.; Bowyer, P.; Matsuda, Y.; Mondo, S.; Lyhne, E. K.; Kogle, M. E.; Clum, A.; Lipzen, A.; Salamov, A.; Ngan, C. Y.; Daum, C.; Chiniquy, J.; Barry, K.; LaButti, K.; Haridas, S.; Simmons, B. A.; Magnuson, J. K.; Mortensen, U. H.; Larsen, T. O.; Grigoriev, I. V.; Baker, S. E.; Andersen, M. R., *Proc. Natl. Acad. Sci. U. S. A.* **2018,** *115*, E753-E761; b) Kjærbølling, I.; Vesth, T.; Frisvad, J. C.; Nybo, J. L.; Theobald, S.; Kildgaard, S.; Petersen, T. I.; Kuo, A.; Sato, A.; Lyhne, E. K.; Kogle, M. E.; Wiebenga, A.; Kun, R. S.; Lubbers, R. J. M.; Mäkelä, M. R.; Barry, K.; Chovatia, M.; Clum, A.; Daum, C.; Haridas, S.; He, G.; LaButti, K.; Lipzen, A.; Mondo, S.; Pangilinan, J.; Riley, R.; Salamov, A.; Simmons, B. A.; Magnuson, J. K.; Henrissat, B.; Mortensen, U. H.; Larsen, T. O.; de Vries, R. P.; Grigoriev, I. V.; Machida, M.; Baker, S. E.; Andersen, M. R., *Nat. Commun.* **2020,** *11*, 1106; c) Gilchrist, C. L. M.; Lacey, H. J.; Vuong, D.; Pitt, J. I.; Lange, L.; Lacey, E.; Pilgaard, B.; Chooi, Y.-H.; Piggott, A. M., *Fungal Genet. Biol.* **2020,** *143*, 103435.

10. Chooi, Y.-H.; Fang, J.; Liu, H.; Filler, S. G.; Wang, P.; Tang, Y., *Org. Lett.* **2013,** *15*, 780-783.

11. Yin, W.-B.; Chooi, Y. H.; Smith, A. R.; Cacho, R. A.; Hu, Y.; White, T. C.; Tang, Y., *ACS Synth. Biol.* **2013,** *2*, 629-634.

12. a) Pitt, J. I.; Lange, L.; Lacey, A. E.; Vuong, D.; Midgley, D. J.; Greenfield, P.; Bradbury, M. I.; Lacey, E.; Busk, P. K.; Pilgaard, B.; Chooi, Y.-H.; Piggott, A. M., *PLoS One* **2017,** *12*, e0170254; b) Li, H.; Lacey, A. E.; Shu, S.; Kalaitzis, J. A.; Vuong, D.; Crombie, A.; Hu, J.; Gilchrist, C. L. M.; Lacey, E.; Piggott, A. M.; Chooi, Y.-H., *Org. Biomol. Chem.* **2021,** *19*, 587-595.

13. a) Bergmann, S.; Schumann, J.; Scherlach, K.; Lange, C.; Brakhage, A. A.; Hertweck, C., *Nat. Chem. Biol.* **2007,** *3*, 213-217; b) Li, H.; Wei, H.; Hu, J.; Lacey, E.; Sobolev, A. N.; Stubbs, K. A.; Solomon, P. S.; Chooi, Y.-H., *ACS Chem. Biol.* **2020,** *15*, 226-233.

14. König, C. C.; Scherlach, K.; Schroeckh, V.; Horn, F.; Nietzsche, S.; Brakhage, A. A.; Hertweck, C., *ChemBioChem* **2013,** *14*, 938-942.

15. Chiang, Y.-M.; Ahuja, M.; Oakley, C. E.; Entwistle, R.; Asokan, A.; Zutz, C.; Wang, C. C. C.; Oakley, B. R., *Angew. Chem. Int. Ed.* **2016,** *55*, 1662-1665.

16. a) Hu, J.; Sarrami, F.; Li, H.; Zhang, G.; Stubbs, K. A.; Lacey, E.; Stewart, S. G.; Karton, A.; Piggott, Andrew M.; Chooi, Y.-H., *Chem. Sci.* **2019,** *10*, 1457-1465; b) Li, H.; Hu, J.; Wei, H.; Solomon, P. S.; Vuong, D.; Lacey, E.; Stubbs, K. A.; Piggott, A. M.; Chooi, Y.-H., *Org. Lett.* **2018,** *20*, 6148-6152.

17. Song, Z.; Bakeer, W.; Marshall, J. W.; Yakasai, A. A.; Khalid, R. M.; Collemare, J.; Skellam, E.; Tharreau, D.; Lebrun, M.-H.; Lazarus, C. M.; Bailey, A. M.; Simpson, T. J.; Cox, R. J., *Chem. Sci.* **2015,** *6*, 4837-4845.

18. Harvey, C. J. B.; Tang, M.; Schlecht, U.; Horecka, J.; Fischer, C. R.; Lin, H.-C.; Li, J.; Naughton, B.; Cherry, J.; Miranda, M.; Li, Y. F.; Chu, A. M.; Hennessy, J. R.; Vandova, G. A.; Inglis, D.; Aiyar, R. S.; Steinmetz, L. M.; Davis, R. W.; Medema, M. H.; Sattely, E.; Khosla, C.; Onge, R. P. S.; Tang, Y.; Hillenmeyer, M. E., *Sci. Adv.* **2018,** *4*, eaar5459.

19. Shen, M. W. Y.; Fang, F.; Sandmeyer, S.; Da Silva, N. A., *Yeast* **2012,** *29*, 495-503.

20. a) Ugai, T.; Minami, A.; Fujii, R.; Tanaka, M.; Oguri, H.; Gomi, K.; Oikawa, H., *Chem. Commun.* **2015,** *51*, 1878-1881; b) Li, H.; Hu, J.; Wei, H.; Solomon, P. S.; Stubbs, K. A.; Chooi, Y.-H., *Chem. Eur. J.* **2019,** *25*, 15062-15066; c) Hewage, R. T.; Huang, R.-J.; Lai, S.-J.; Lien, Y.-C.; Weng, S.-H.; Li, D.; Chen, Y.-J.; Wu, S.-H.; Chein, R.-J.; Lin, H.-C., *Org. Lett.* **2021,** *23*, 1904-1909.

21. Gilchrist, C. L. M.; Booth, T. J.; van Wersch, B.; van Grieken, L.; Medema, M. H.; Chooi, Y.-H., *Bioinformatics Advances* **2021,** *1*, vbab016.