Heterologous Biosynthesis and Genomics-Driven Derivatization of Fungal Bioactive Sesterterpenoid Variecolin

Dexiu Yan,† Jemma Arakelyan,† Teng Wan,† Tsz Ki Chan,‡ Dohyun Ahn,† Tsz Kiu Cheung,† Ho Ching Chan,† Inseo Choi,§ Pui Yi Ho,† Feijun Hu,† Yujeong Kim,† Hill Lam Lau,† Ying Lo Law,† Chi Seng Leung,† Ritvik Raina,‡ Chun Yin Tong,† Kai Kap Wong,† Wing Lam Yim,† Richard Y. C. Kong,*† Maria V. Babak,*† and Yudai Matsuda*‡

†Department of Chemistry, City University of Hong Kong, Tat Chee Avenue, Kowloon, Hong Kong SAR, China
‡Department of Biomedical Engineering, City University of Hong Kong, Tat Chee Avenue, Kowloon, Hong Kong SAR, China

ABSTRACT: The biosynthetic gene cluster of fungal bioactive sesterterpenoids, variecolin (1) and variecolactone (2), was identified in Aspergillus aculeatus ATCC 16872. Heterologous production of 1 and 2 was achieved in Aspergillus oryzae by expressing the sesterterpene synthase VrcA and the cytochrome P450 VrcB. Intriguingly, the replacement of VrcB with homologous P450s from other fungal terpenoid pathways yielded three new variecolin analogues, one of which exhibited potent anticancer activity comparable to that of 1.

Sesterterpenoids, derived from five molecules of isoprene (C₅) units, are one of the classes of terpenoids, which are widely distributed but relatively rare in nature.¹ Fungi are prolific producers of sesterterpenoids with a wide range of biological activities, as represented by ophiobolin A that displays a potent anticancer activity.² Since the discovery and characterization of the first sesterterpene synthase in 2013,³ biosynthetic pathways of several fungal sesterterpenoids have been elucidated to date.⁴ Furthermore, a number of new sesterterpenoids have been obtained by genome mining approach targeting unexploited fungal sesterterpene synthase genes.⁵ All of the characterized fungal sesterterpene synthases are chimeric proteins consisting of C-terminal prenyltransferase (PT) and N-terminal terpene cyclase (TC) domains. The PT domain is responsible for the synthesis of C₁₅ geranylgeranyl pyrophosphate (GGPP) using one molecule of dimethylallyl pyrophosphate (DMAPP) and four molecules of isopentenyl pyrophosphate (IPP) as substrates, whereas the TC domain performs ionization-initiated (class I) cyclization of GGPP to afford a cyclized product.⁶ The cyclized product could then undergo tailoring reactions to be converted into a final sesterterpenoid. Interestingly, the biosynthetic gene clusters of fungal sesterterpenoids are relatively compact compared with those for other fungal secondary metabolites, suggesting that fungal sesterterpenoid biosynthesis could easily be engineered to generate diverse analogues.

Variecolin (1) is a tetracyclic sesterterpenoid originally isolated from the fungus Aspergillus variecolor MF138 as an angiotensin II receptor binding inhibitor.⁷ Subsequently, 1 and its congeners, such as variecolactone (2) and emericolin A (3), have also been obtained from several different fungi and have proven to possess diverse and potent biological activities (Figure 1A),⁸ including antifungal,⁹ immunomodulatory,¹⁰ CCRS inhibitory,¹¹ antibacterial,¹² antimalarial,¹³ and anticancer¹⁴ activities. The biological activities and the intriguing molecular framework of 1 have also inspired synthetic chemists to attempt the synthesis of 1, however, no total synthesis of 1 has been achieved, making 1 a challenging synthetic target. Furthermore, despite the recent rapid advance in biosynthetic studies on fungal sesterterpenoids, the biosynthetic genes for 1 have not been reported. Thus, identification and characterization of the biosynthetic gene cluster of 1 would facilitate efficient production and structural modification of this medicinally important molecule in biocatalytic or synthetic biology approaches.

In this study, to achieve the total biosynthesis and derivatization of variecolin (1), we first identified the biosynthetic gene cluster of 1 in Aspergillus aculeatus, and then succeeded in the heterologous production of 1, which requires the terpene synthase VrcA and the cytochrome P450 monoxygenase VrcB. We also obtained three new variecolin analogues by utilizing a series of VrcB homologues. Finally, we evaluated the biological activity of the obtained molecules, and found that one of the unnatural variecolin analogues displays potent anticancer activity.
and of the transformant, formant harboring to a sesterterpene hydrocarbon, was detected. As a result, a new product reconstitution of fungal metabolite biosynthesis. The investigation of two homologous proteins of known sesterterpene synthases in fungi, including another known producer of variecolin biosynthesis, hereby designated as the terpene synthase gene vrcA. Further analysis of 2D NMR spectra illuminated that the enzyme is involved in the variecolin pathway. Thus, the terpene synthase gene vrcA+Stl-P450, which is followed by the cyclization of GFPP to form 4 should occur in a similar manner to that of astellifadiene (Figure 3), as previously discussed. Thus, after the departure of the pyrophosphate group, C-C bond formations occur at C-1/C-11 and C-10/C-14 to generate the bicyclic tricyclic cation 4b+ with 11-5-fused ring system. Next, the five-membered ring undergoes ring expansion, which is concerted with the additional ring formation at C-14/C-18. The resultant 11-6-5-fused ring system is then neutralized by the deprotonation from C-24 to yield 4c. The absolute configuration of 4 was deduced based on that of 1, and 4 was named variecoladiene (Figure 2C).

Figure 1. (A) Structures of variecolin (1) and its selected analogues. (B) Schematic representation of the vrc cluster. (C) Predicted biosynthetic pathway of 1.

To discover the biosynthetic gene cluster of variecolin (1), we examined the genome sequence of A. aculeatus ATCC 16872 (CBS 172.66), since A. aculeatus is one of the known producers of 1. The investigation of the A. aculeatus genome revealed the presence of two homologous proteins of known sesterterpene synthases. Since the backbone skeleton of 1 is expected to be synthesized in a similar manner to the generation of astellifadiene, the terpene synthase for the variecolin biosynthesis should display high similarity to the generation of astellifadiene (Figure 1B). The gene cluster, which is hereby designated as the vrc cluster, is conserved among several fungi, including another known producer of 1, Aspergillus japonicus, further indicating the involvement of the vrc cluster in the variecolin biosynthesis. On the basis of the gene cluster information, the biosynthetic route leading to 1 could be predicted as follows; the sesterterpene synthase VrcA first yields a hydrocarbon with the variecolin scaffold, which then undergoes multiple oxidations catalyzed by the P450 VrcB to afford 1 (Figure 1C).

To analyze the function of the vrc cluster, we next heterologously expressed the terpene synthase gene vrcA in the fungus Aspergillus oryzae NSAR1, which is a powerful platform for the heterologous reconstitution of fungal metabolite biosynthesis. The metabolites derived from the A. oryzae transformant were analyzed by GC-MS. As a result, a new product 4 with m/z 340 [M]+, which corresponds to a sesterterpene hydrocarbon, was detected only in the transformant harboring vrcA (Figure 2A). After the large-scale cultivation of the transformant, 4 was successfully isolated and subjected to MS and NMR analysis for structural characterization. The HR-APCI-MS analysis determined the molecular formula of 4 as C_{34}H_{30} with six degrees of unsaturation. Four olefinic carbon signals were observed in the 13C NMR spectrum of 4, collectively indicating the tetracyclic nature of 4. Further analysis of 2D NMR spectra illuminated that 4 possesses the identical carbon skeleton to that of variecolin (1) (Figure S1), and therefore, 4 is likely to be the biosynthetic precursor of 1. The absolute configuration of 4 was deduced based on that of 1, and 4 was named variecoladiene (Figure 2C).
We then co-expressed the P450 gene vrcB along with the terpene synthase gene vrcA in A. oryzae, and consequently, two major peaks were observed in the HPLC analysis of the metabolites produced by the two gene-expressing strain (Figure 2B, trace iii). These two metabolites were purified, and the NMR analyses elucidated that they are variecolin (1) and variecolactone (2), respectively. In the course of the purification procedure, we noticed the presence of another minor metabolite, which was also isolated and identified as emericolin A (3). Thus, it is indicated that the P450 VrcB is solely responsible for the multiple oxidations at C-5 and C-20 positions. Collectively, it has been proven that the vrc cluster is indeed responsible for the variecolin biosynthesis, and heterologous production of 1 was successfully achieved.

Having established the heterologous production platform of variecolin (1), we then sought to derivatize the variecolin scaffold using the A. oryzae system. To this end, we collected known P450 genes involved in fungal di-/sesquerolpenoid biosynthesis as well as uncharacterized P450 genes that are clustered with a potential sesquerolpenyl synthase gene (Table S2 and Figures S2 and S3), and used them as genetic building blocks to construct artificial sesquerolpenoid pathways. Each P450 gene was individually co-expressed with the sesquerolpenyl synthase gene vrcA, and among the 17 tested P450s, four P450s successfully yielded oxidized products of the hydrocarbon 4. Two P450s from Aspergillus violaceofuscus CBS 1157114 (designated as VrcB; GenBank: PY14930.1) and Aspergillus heteromorphus CBS 117.5514 (designated as VrcB; GenBank: XP_025395330.1) exhibited the same product profile as VrcB (Figure 2B, trace iv and v), and this observation would be reasonable, given the high sequence similarity between VrcB and the two P450s (89% and 79% amino acid identity, respectively). Intriguingly, the introduction of acldA-P450 involved in the biosynthesis of asperterpenol B14 afforded a new metabolite 5 (Figure 2B, trace vi), which was found to be a monooxygenated form of 4 by HR-MS analysis. The 1H NMR spectrum of 5 revealed the presence of an oxymethine signal (δH 4.01 ppm), and accordingly, a signal for an oxygenated carbon (δC 79.9 ppm) was observed in the 13C NMR spectrum, altogether indicating that 5 is a hydroxylated form of 4.

Further analysis of the 2D NMR spectra confirmed that 5 harbors the α-hydroxy group at C-4 (Figures 2C and S1), as observed in asperterpenol B. Meanwhile, Stl-P450 responsible for the stellatic acid biosynthesis24 transformed 4 to two products 6 and 7 (Figure 2B, trace vii), both of which are isomers of 5. The NMR spectra of 5 and 6 highly resemble each other, and the signals for the exomethylene at C-23/C-24 disappeared in both compounds. Alternatively, two doublet signals were newly observed at 2.5–2.7 ppm in the 1H NMR spectra, and likewise, two signals at 5.5–6.1 ppm appeared in the 13C NMR spectra, implying that the exomethylene was epoxidized by Stl-P450. The interpretation of the 2D NMR spectra then confirmed that 6 and 7 have the same planar structure with the epoxide at C-23/C-24 and only differ at the stereochemistry of C-23 position. The stereochemistry of the epoxide in 6 and 7 was determined to be R and S, respectively, based on the NOESY correlations (Figures 2C and S1). It should be noted that the 1H signal for H-9a of 7 (δH 2.58 ppm) was observed at a much lower magnetic field than that of 6 (δH 2.18 ppm). This observation indicates that the epoxide oxygen of 7 is more proximately located to H-9a than in 6, further supporting the deduced stereochemistry of the two compounds (Figure S1). Interestingly, Stl-P450 performs the oxidation of a methyl group to a carboxy group in the stellatic acid biosynthesis, which is distinct from the reaction that occurred on the variecolin scaffold.

Considering that variecolin (1) and its analogues exhibit diverse biological activities, it was expected that new metabolites obtained in this study could also have similar activities. All compounds, except for 3, were tested for cytotoxicity against human breast adenocarcinoma cell line MCF-7 and murine colon cancer cell line CT26 (Table 1 and Figures S4 and S5). The hydrocarbon 4 demonstrated only marginal cytotoxicity in the high micromolar concentration range, whereas the other compounds demonstrated improved anticancer activity in both cell lines. Remarkably, compound 5 exhibited comparable cytotoxicity to variecolin.

**Table 1. Cytotoxicity of the metabolites obtained in this study determined by colorimetric MTT assay.**

<table>
<thead>
<tr>
<th>Compound</th>
<th>MCF-7 [μM]</th>
<th>CT26 [μM]</th>
</tr>
</thead>
<tbody>
<tr>
<td>variecolin (1)</td>
<td>1.4 ± 0.3</td>
<td>1.0 ± 0.3</td>
</tr>
<tr>
<td>variecolactone (2)</td>
<td>22 ± 8</td>
<td>14 ± 3</td>
</tr>
<tr>
<td>variecoladiene (4)</td>
<td>&gt;500</td>
<td>409 ± 131</td>
</tr>
<tr>
<td>5</td>
<td>8 ± 3</td>
<td>2.9 ± 0.4</td>
</tr>
<tr>
<td>6</td>
<td>146 ± 82</td>
<td>44 ± 8</td>
</tr>
<tr>
<td>7</td>
<td>31 ± 9</td>
<td>16 ± 5</td>
</tr>
</tbody>
</table>

In conclusion, in this study, we have successfully achieved the heterologous production of variecolin (1), which has long been a challenging target for synthetic chemists. Furthermore, we have derivatized the variecolin scaffold by employing P450s from other fungal terpenoid pathways, and thus obtained three distinct biocatalysts that can efficiently oxidize the variecolin skeleton. Given that new variecolin analogues with different oxidative modifications all
exhibited anticancer activity, the variecolin skeleton could be considered as a privileged scaffold for developing anticancer agents. Our study has demonstrated the usefulness of our simple synthetic biology strategy to obtain "unnatural" natural products. This approach could also be applied for the structural modification of other fungal sesterterpenoids, which would further diversify this class of natural products and possibly provide future drug leads with improved biological activities.

**ASSOCIATED CONTENT**

**Supporting Information**
A Supporting Information file is associated with this manuscript.

**AUTHOR INFORMATION**

**Corresponding Author**
*bhrkong@cityu.edu.hk (R.Y.C.K.)
*mbabak@cityu.edu.hk (M.Y.B.)
ymatsu@cityu.edu.hk (Y.M.)

**Notes**

The authors declare no competing interest.

**ACKNOWLEDGMENT**

This work was performed for the iGEM 2021 Competition; financial support from the Department of Chemistry, City University of Hong Kong for the competition is greatly appreciated. We thank Prof. Katsuya Gomi (Tohoku University) and Profs. Katsuhiko Kitamoto and Jun-ichi Maruyama (The University of Tokyo) for the expression vectors and the fungal strain. We are grateful to Dr. Man-Kit Tse (City University of Hong Kong) for his assistance in NMR spectra acquisition. This work was in part supported by the Early Career Scheme grant from the Research Grants Council (RGC) of Hong Kong (Project No. 21300219 (Y.M.)).

**REFERENCES**


