Synthetic biology based construction of biological activity-related library of fungal decalin-containing diterpenoid pyrones

Kento Tsukada1, Shono Shinki1, Akiho Kaneko1, Kazuma Murakami2, Kazuhiro Irie2, Masatoshi Murai3, Hideto Miyoshi3, Shingo Dan4, Kumi Kawaji5, Hironori Hayashi6, Eiichi N. Kodama5,6, Aki Hori7, Emil Salim7, Takayuki Kuraishi7, Naoya Hirata8, Yasunari Kanda8 & Teigo Asai1,9

A synthetic biology method based on heterologous biosynthesis coupled with genome mining is a promising approach for increasing the opportunities to rationally access natural product with novel structures and biological activities through total biosynthesis and combinatorial biosynthesis. Here, we demonstrate the advantage of the synthetic biology method to explore biological activity-related chemical space through the comprehensive heterologous biosynthesis of fungal decalin-containing diterpenoid pyrones (DDPs). Genome mining reveals putative DDP biosynthetic gene clusters distributed in five fungal genera. In addition, we design extended DDP pathways by combinatorial biosynthesis. In total, ten DDP pathways, including five native pathways, four extended pathways and one shunt pathway, are heterologously reconstituted in a genetically tractable heterologous host, Aspergillus oryzae, resulting in the production of 22 DDPs, including 15 new analogues. We also demonstrate the advantage of expanding the diversity of DDPs to probe various bioactive molecules through a wide range of biological evaluations.
N
atural products are historically unparalleled successful sources for drug discovery1–3. During secondary metabolism biosynthesis in living organisms, natural products are formed in vivo through multi-step enzymatic reactions4,5. In other words, natural product structures are constructed and tailored in repeated interaction with proteins in cells. This biosynthetic manner provides natural product with drug-advantageous properties such as potential protein interactivity, water solubility, and membrane permeability6. For this reason, the natural product chemical space is highly relevant to biological space7. Therefore, the discovery of natural products with novel structures and biological activities is an important challenge in the pharmaceutical field. Natural products with valuable biological activities are recognized to contain biologically relevant privileged structures,7,8 whose analogues and congeners serve as excellent guides to discover not only more potent bioactive compounds but also various additional biological activities9–10. Thus, expanding the chemical space around biologically important natural products may accelerate drug discovery research11. Chemical synthesis, including divergent and diverted total syntheses, is a powerful strategy to produce bioactive natural product analogues and congeners12,13. A late-stage modification strategy is also a beneficial tool to generate natural product derivatives, such as diversity-oriented synthesis using natural products and their intermediate starting points14–17. However, the structural complexity and limited availability of natural products remain obstacles to synthesizing a large collection of natural products and their structural analogues in sufficient amounts. Thus, the development of a method that enables rapid and rational access to the chemical space around bioactive natural products and a reliable supply is required for drug discovery.

Fungi are among the most important microbial resources for drug discovery because of their ability to produce structurally diverse and biologically important natural products18. It is also well known that fungi possess extraordinary biosynthetic gene clusters that may encode highly diverse natural product biosynthetic pathways, and biosynthetic gene information has been accumulating at a rapidly accelerating rate in the past a decade because of genome sequencing innovation19. However, most biosynthetic pathways distributed in fungal genomes have not been linked with structural information20. This implies the presence of a large number of novel natural products that may be produced via those untapped pathways. A synthetic biology method based on heterologous biosynthesis coupled with genome mining is a promising approach to translate enormous amounts of biosynthetic gene information to richly diverse natural products. Genome mining based on biosynthetic studies has enabled rapid access to biosynthetic pathways not only for a target natural product and its related analogues but also for unexplored natural products. Heterologous biosynthesis has enabled to access compounds encoded by biosynthetic pathways found by genome mining and a number of total biosynthesis of natural products and discovering novel natural products have been reported21–28. In addition, combinatorial biosynthesis approach is a powerful tool to generate non-natural analogues, which may be a great advantage to construct pharmaceutical beneficial screening library29–32. Thus, we apply the synthetic biology approach for rationally expanding the chemical space around a target natural product as follows (Fig. 1). The genome mining and reconstruction of a biosynthetic pathway for a target natural product in a heterologous host achieves its total biosynthesis, and the biosynthetic information allows us to mine its related biosynthetic pathways and access its analogues. We, then, apply pathway extension for combinatorial biosynthesis that can produce new natural product analogues that are not programmed in nature and more highly modified than programmed natural products. Fungi produce decalin-containing diterpenoid pyrones (DDPs), a type of meroterpenoid natural products composed of a diterpenoid-derived decalin ring system linked to pyrone biosynthesized via polyketide pathways33. DDPs are found only in fungi, and over 20 DDPs have been reported to date (Supplementary Fig. 1). Interestingly, even small structural differences in each DDP show a wide range of biological activities, such as antiproliferative activity against cancer cell lines and immunosuppressive activity, implying that DDPs contain a privileged structure (Fig. 2)33–41. Thus, the chemical space of DDPs is likely relevant to a diverse biological space and expanding this space leads to the development of a valuable screening library for drug discovery. In addition to promising biological activities, the structural features of DDPs render them exciting targets for total synthesis42–44. Recently, excellent divergent total synthesis has been reported, which enables access to four DDPs via a common intermediate in less than twenty chemical reaction steps45. Nevertheless, the synthesis of diverse DDPs as well as non-natural analogues is a rather elaborate effort; and therefore, it is difficult to prepare a large collection of DDPs and expand their chemical space by chemical means.

Here, we demonstrate the advantage of the synthetic biology approach based on heterologous and combinatorial biosynthesis coupled with genome mining for constructing a biologically relevant DDP-focused library composed of a variety of DDPs. Although biosynthetic studies on DDPs are limited, a putative biosynthetic gene cluster and pathway for subglutinols (subA–F) in an entomopathogenic fungus, Metarhizium robertsi, has been estimated (Fig. 3a). In addition, the functions of a non-reducing polyketide synthase (NR-PKS) SubA, geranylgeranyl diphosphate synthase (GGPPS) SubD and prenyltransferase (PT) SubC have been identified by a heterologous expression study46. We perform genome mining based on this information and find a putative DDP gene clusters distributed in five fungal genera Arthrinium,
Metarhizium, Colletotrichum, Macrophomina and Fusarium fungi (Fig. 3b). According to bioinformatics analyses, we design five native pathways from those biosynthetic gene clusters and reconstruct them in Aspergillus oryzae NSAR122–27, 47, an excellent heterologous host for the production of fungal natural products, to give intermediates and end products encoded in all the pathways. Subsequently, we conduct pathway extension for combinatorial biosynthesis by adding additional modification enzymes to the native DDP pathways, yielding unnatural DDP analogues. Overall, we successfully produce 22 DDPs including 15 analogues that have not been reported, which include intermediates, end products and additionally modified analogues. Because they all can be easily re-supplied by cultivation of the corresponding transformant, we are able to evaluate a variety of biological activities of the DDP-focused library and find wide range of potent bioactivities, such as cell cytotoxicity against cancer cell lines through mitochondrial complex III inhibition, antiproliferative activity against cancer stem-like cells, anti-HIV, preventing amyloid 3(4)-amyloid aggregation in nucleation phase, paralysing activity against adult Drosophila and suppressing insect innate immune signal transduction. Most these biological activities firstly have been found in DDPs in this study.

Results

Genome mining and design of DDP biosynthetic pathways. To find biosynthetic gene clusters that may encode DDP pathways, we perform genome mining of the public databases and our original gene resources by using SubA (NR-PKS) as a query. As a result, five candidate gene clusters with subA-E orthologous genes that may encode NR-PKS, GGPPS, PT, flavin adenine dinucleotide (FAD)-dependent epoxidase (FMOep) and terpene cyclase (TC) were found in five fungal genomes, Fusarium graminearum PH-1 (dpfgABCDEFGHIIJK), Macrophomina phaseolina MS6 (dpmpABCDEFGHIIJ), Colletotrichum higginsianum IMI349063 (dpchABCDEFGH), Metarhizium anisopliae E6 (dpmaABCDEF), and Arthrinum sacchari (dpasABCDFE), which we previously isolated from a spider (Fig. 3b and Supplementary Table 1). The dpma gene cluster, which is identical to the sub gene cluster in M. robertsii, is widely conserved across the genus Metarhizium, and the dpfg gene cluster is also broadly distributed in the genus Fusarium (Supplementary Fig. 2). To design native DDP biosynthetic pathways distributed in the five fungal genera, the five gene clusters were comparatively analysed based on amino acid sequence homology and reordered them as shown in Fig. 3c (Supplementary Tables 2 and 3). The five genes (dpfAEPKDE) are highly conserved in each cluster, suggesting that all the pathways share the biosynthetic pathway for a common intermediate 4. The differences in the genes at the tailoring steps in each pathway may diversify DDP biosynthesis. The three biosynthetic gene clusters in F. graminearum (dpf cluster), M. phaseolina (dpma cluster) and C. higginsianum (dpch cluster) contain two types of short chain dehydrogenase reductase (SDR) genes, dpfXG (SDR1) and dpfXH (SDR2). Each of the dpfXG and dpfXH genes can be recognized as orthologous because of their high similarity (their encoded enzymes are approximately 60% identical each other); that is, each set of SDRs, dpfGH, dpmpGH and dpchGH, may produce the same product from 4. Both the dpf and dpmp clusters include an orthologous methyltransferase gene, dpfXl (MT1), which it is absent in the dpch cluster, indicating that the dpf pathway is probably divided from the dpf and dpmp pathways after SDR modification steps and that a remaining dpchF (FMO) would lead to the end product in the dpch pathway. Both the dpmp and dpf clusters contain a P450 gene, dpmpJ and dpfJ, respectively, but they show low similarity to each other (their encoded enzymes are ~20% identical to each other), suggesting that the dpf and dpmp pathways may have branched after the MT1 modification step. Subsequently, dpmpJ and dpfJK afford the final products in each pathway. However, the dpas and dpma gene clusters possess only a modifying enzyme, FMO dpasF and FAD-dependent BBE domain-containing oxidoreductase (BBE) dpmaF, respectively. Therefore, both pathways are likely branched at the initial tailoring stage, leading to the final products. As a result, we predicted the treelike native DDP pathways, as depicted in Fig. 3d. Considering previously reported DDPs and their producing fungi, the dpch and dpma pathways may produce higginsianin A40 and subglutinol A8, 40, respectively. However, the orphan dpfg, dpmp and dpas pathways may provide new DDP analogues.

Reconstitution of DDP pathways. We reconstructed all the pathways by stepwise introduction of the biosynthetic genes in a heterologous host A. oryzae NSAR1 according to the hypothetical biosynthetic pathways (Fig. 3d and Supplementary Fig. 5). Because it is difficult to obtain genome-sequenced strains of F. graminearum PH-1, M. phaseolina MS6, C. higginsianum IMI349063 and M. anisopliae E6, we used F. graminearum 50218, M. phaseolina NBRRC7317, C. higginsianum MAFF305635 and M. anisopliae NBRCC103233 instead as gene donors.

Initially, we introduced dpasACD into A. oryzae to construct the A. oryzae transformant with dpasACD (AO-dpasACD), which, as expected, produced prenylated (C20) α-pyrene 2 (Fig. 3a and Supplementary Fig. 7). We subsequently introduced dpasBE into AO-dpasACD to construct AO-dpasABCDE, which provided the common intermediate 4 (87 mg L−1) (Supplementary Fig. 7), demonstrating its biosynthetic machinery for the first time. We also characterized dpfBE, dpmpBE or dpchBE as sharing the same function of dpasBE by introducing them into AO-dpasACD (Supplementary Fig. 7).
**Fig. 3** Genome mining and design of DDP pathways. 

**a** Biosynthetic pathway for common intermediate 4. 

**b** DDP biosynthetic gene clusters distributed in five fungal genera. 

**c** Comparative analysis of each biosynthetic gene cluster. 

**d** Design of native pathways, extended steps and one shunt pathway heterologously reconstituted in this study and summary of products (red numbers show compounds that have not been reported.) produced through the DDP pathways (blue square, green square, red square and black square show intermediates, end products, additionally modified products and shunt products, respectively).
Next, we aimed to comprehensively reconstitute modification steps in native DDP pathways by using the 4-producing transformants as a platform. We reconstituted the dpas and dpma pathways by introducing an FMO gene \( \text{dpasF} \) or a BBE gene \( \text{dpmaF} \) into \( \text{AO-dpasABCDE} \) to give \( \text{AO-dpasABCDEF} \) and \( \text{AO-dpasABCDE-dpmaF} \). \( \text{AO-dpasABCDE-dpmaF} \) provided subglutinols A (5, 62 mg L\(^{-1}\)) and B (6, 5 mg L\(^{-1}\)) (Supplementary Figs. 12 and 13). On the other hand, \( \text{AO-dpasABCDEF} \) produced a new DDP analogue with an enone system at the C5 unit (7, 6 mg L\(^{-1}\)) as well as subglutinols A (5, 62 mg L\(^{-1}\)) and B (6, 5 mg L\(^{-1}\)) (Supplementary Figs. 12 and 13). These results suggested that both DpasF and DpmaF are involved in tetrahydrofuran (THF) ring formation at the C5 unit, while DpasF possesses an additional catalytic ability of multi-step oxidations to generate the enone at the C5 unit (Fig. 4b).

We then aimed to reconstitute the dpch, dpmp and dpfg pathways in \( \text{A. oryzae} \). The introduction of \( \text{dpmpG} \) or \( \text{dpmpGH} \) into \( \text{AO-dpasABCDE} \) afforded two transformants, \( \text{AO-dpasABCDE-dpmpG} \) and \( \text{AO-dpasABCDE-dpmpGH} \). AO-dpasABCDE-dpmpG
accumulated ketone 8 (23 mg L\(^{-1}\)), while AO-dpasABCDE-dpmpGH gave higginsinian B 9 (28 mg L\(^{-1}\)) (Supplementary Fig. 17). This result showed that SDR1, DpmpG, oxidized the 8S hydroxy group to a ketone and SDR2, DpmpH, reduced the ketone to the 8 R hydroxy group in a similar manner to that in andrastin biosynthesis.\(^{48}\) (Fig. 4a). We also demonstrated that dpfglGH and dpchGH were involved in the same inversion of the stereocchemistry at C-8 (Supplementary Fig. 17). The entire dpch pathway was heterologously reconstituted by introducing dpchGH into a 4-producing transformant, and the resulting transformant, as expected, produced higginsinian A 10 (28 mg L\(^{-1}\)) (Fig. 4b and Supplementary Fig. 21). We reconstituted whole dpmp and dpfg pathways by using higginsinian B 9 producing transformants as a platform. The introduction of dpmpI and dpfgl into the platforms yielded AO-dpasABCDE-dpmpGH1 and AO-dpasABCDE-dpfgGHI, both of which produced a new intermediate 11 (29 mg L\(^{-1}\) and 20 mg L\(^{-1}\), respectively). Moreover, dpmpI and dpfgl were introduced to construct AO-dpasABCDE-dpmpGHI1 and AO-dpasABCDE-dpfgGHIJK, the transformant expressing the whole dpmp gene cluster, AO-dpasABCDE-dpmpGHI, afforded a new DDP with a C-26 primary alcohol on the pyrone moiety (12, 21 mg L\(^{-1}\)) (Fig. 4c and Supplementary Fig. 25). On the other hand, the transformant expressing the whole dpfg pathway, AO-dpasABCDE-dpfgGHIJK, produced the new DDPs 13 (major product, 5 mg L\(^{-1}\)) and 14 (minor product, 2 mg L\(^{-1}\)) with a highly oxidized y-pyrene moiety including a methyl ester (Fig. 4c and Supplementary Fig. 25), like colletotrichin obtained from C. nicotianae.\(^{41}\) The results suggested that a P450, DpmpI, oxidized C-26 methyl to primary alcohol, Dpfgl, catalysed a three-step oxidation at C-27 to generate a carboxylic acid as well as C-26 hydroxylation. The results also indicated that an MT1, Dpfgl, is involved in the same methylation as DpmpI, while an MT2, Dpfgk, methylates the carboxylic acid generated by Dpfgl. We thus completely reconstructed all the DPP native pathways distributed in five fungi, resulting in the production of 11 DDPs, including five new analogues. As expected, the dpma and dpch pathways encoded the biosynthetic pathways for subglutinols 5, 6) and higginsinian A 10, respectively, and the orphan dpas and dpch pathways provided new DDP analogues.

We also investigated substrate selectivities of the modification enzymes using A. oryzae heterologous expression system. In summary, DpmaF and Dpfgl strictly recognized the C-8 configuration, while DpasF, DpchF, DpmpI, Dpfgl and Dpmpj showed tolerant substrate selectivity, which became an advantage for the generating diversity in combinatorial biosynthesis. Through the experiments, we found and reconstituted a shunt DDP pathway in A. oryzae, which afforded viridoxin A hydrolysate 15 (42 mg L\(^{-1}\)) and a new DDP analogue 16 (6 mg L\(^{-1}\)) (Figs. 3d, 4e, Supplementary Figs. 29, 31 and Supplementary Note 15).

**Pathway extension for combinatorial biosynthesis.** The reactions of each enzyme in all the pathways are summarized in Fig. 4a-c. The enzymes that catalysed C\(_{8}\) unit modifications were specifically distributed in the dpma, dpas and dpch pathways, while the enzymes involved in the pyrene moiety modifications were localized in the other pathways. However, no pathway that containing both C\(_{8}\) unit and pyrene moiety modification enzymes is encoded in native DDP biosynthetic gene clusters. Therefore, we aimed to generate unnatural DDPs with further modified structures than those of the end products in each DDP pathway and conducted combinatorial biosynthesis by combining C\(_{8}\) unit-modifying pathways with pyrone-decorating pathways (Fig. 3d). We chose dpasF among C\(_{8}\) unit-modifying enzymes as an additional modification enzyme, because DpasF (FMO) enables the generation of four types of C\(_{8}\) unit moieties through its multi-functional oxidative ability, and we designed two extended pathways by adding dpasF to the dpmp and dpfg pathways. In addition, MT1 (dpmpI) was connected to the dpas and dpch pathways. Thus, we designed and reconstituted the four extended pathways in A. oryzae (Fig. 3d). One transformant (dpmp pathway + dpasF), as expected, produced four new analogues, 17 (5 mg L\(^{-1}\)), 18 (3 mg L\(^{-1}\)), 19 (16 mg L\(^{-1}\)), and 20 (5 mg L\(^{-1}\)), due to the promiscuity of DpmpJ (Figs. 3d, 4d and Supplementary Fig. 35). Another transformant (dpfg pathway + dpasF) afforded only the new analogue 21 (6 mg L\(^{-1}\)), the most modified compound in this study because Dpfgl (P450) strictly recognizes C-8 stereochimistry (Figs. 3d, 4d and Supplementary Fig. 36). In this transformant, another expected product originating from 14 could not be observed in HPLC analysis. The third transformant (dpch pathway + dpmpI), as expected, produced O-methylated higginsinian A 22 (17 mg L\(^{-1}\)), and the fourth transformant (dpas pathway + dpmpI) provided 23 (23 mg L\(^{-1}\)), 24 (7 mg L\(^{-1}\)), and 25 (3 mg L\(^{-1}\)) (Figs. 3d, 4d and Supplementary Fig. 40). As expected, all the non-native analogues produced through combinatorial biosynthesis contained modifications on both the C\(_{8}\) unit and the pyrene moiety. Since the four re-designed extended pathways are not found in the genome databases, of course, every compound produced through the pathways had new structures.

Thus, we achieved the comprehensive production of fungal DDPs through the reconstitution of five-native pathways, one shunt pathway and four extended pathways in A. oryzae and produced 22 DDPs, including 15 new compounds. Among these 22 compounds, 11 compounds came from native pathways, 2 compounds were biosynthesized through shunt pathways, and 9 compounds were produced via extended pathways (Supplementary Table 19). The new compounds, 11-14 and 16-25 were named as FDDP A-O, respectively. All the compounds produced in this study were purified, and their structures were fully determined by spectral analyses. The absolute configuration of the common intermediate 4 was determined by the modified Mosher’s method\(^{49}\) (Supplementary Fig. 10), while those of 5, 6, 9, and 10 were identified by comparing their optical rotation with reported values\(^{40,43}\). The absolute configurations of the other DDPs were determined based on their biosynthetic relationships. The titres of all the DDPs based on HPLC analysis of all the transformants are summarized in Supplementary Table 19.

**Antiproliferative effects on cancer stem-like cells.** Initially, we evaluated new DDP analogues for their antiproliferative activities across the panel of 39 human cancer cell lines, JFCR35\(^{50,51}\). The assay not only showed their cytotoxic effects but also provided the characteristic profiles similar to those of antimycin A and myxothiazol, known inhibitor of mitochondrial complex III (Supplementary Figs. 44, 45 and Supplementary Tables 21, 22). We then revealed that most DDPs, except for enone-containing compounds, selectively prevented mitochondrial complex III, and the effect was the same degree as that of antimycin A in vitro assay (Supplementary Table 20).

We also evaluated antiproliferative activity against cancer stem cells (CSCs), which are a small sub-population in tumour bulk identified in most cancer cells and clinical samples.\(^{52}\) CSCs have been a major problem in cancer therapy because they are responsible for recurrence, metastasis and drug resistance to chemotherapeutic agents that affect proliferative cells.\(^{53-56}\) Anti-CSC activity was evaluated by the inhibition of sphere-forming ability, a well-studied method to enrich the CSC-like population, and selective cytotoxicity against one of the most reliable CSC...
markers, aldehyde dehydrogenase (ALDH)-positive cells. We selected compounds 5, 11, 12, 17, 18, 20, 22 and 23 according to dose-dependent cytotoxicity against the bulk of MCF-7 cells for the mammosphere formation assay (Supplementary Fig. 46a). Low-toxicity 16 was also added as a negative control. Each compound was tested at 5 μM, and 11, 12 and 22 clearly inhibited sphere formation (Fig. 5a, b) more effectively than the anticancer drugs 5-fluorouracil (5-FU) and doxorubicin. Interestingly, 12 showed the strongest effect on sphere formation, while its C-8 epimer 16 did not affect sphere formation, suggesting that the 8R configuration significantly contributed to this activity. We next evaluated the three DDPs that inhibited mammosphere formations against ALDH-positive cells, a functional marker for breast CSCs. MCF-7 cells were treated with 1 μM of 11, 12 or 22, and the ALDH-positive cell population was analysed by using a fluorescence activated cell sorter (Fig. 5c). The results showed that 11, 12 and 22 reduced the population of ALDH-positive cells. The most effective compound, 12, was tested at lower concentrations and showed selective and potent cytotoxicity against ALDH-positive cells (Fig. 5d). Thus, for the first time, we successfully found DDPs with potent antiproliferative effects on the CSC-like population in MCF-7 cells, which may lead to new anti-CSC drugs. Interestingly, subtle structural differences in DDPs generated the difference in anti-CSC activity.

**Biological activity against Drosophila.** We screened the DDPs on the Drosophila assay system for cytotoxicity, inhibitory activity of innate immune signalling and insecticidal activity. The cytotoxicity of the DDPs was evaluated using two cell lines, embryonic macrophage-derived DL1 cells and larval blood cell-derived l(2)mbn cells, and their IC50 values are listed in Supplementary Table 23. Most DDPs except for compounds with the enone at the C5 unit (7, 19 and 25) showed potent antiproliferative activity, and they affected DL1 cell lines more potently. Notably, subglutinol A (5) and 21 strongly inhibited cell growth in the DL-1 cell line (IC50 = 12 nM (5) and 15 nM (21)).

We next tested the ability of the DDPs to affect the innate immune pathways, namely, the Toll and immune deficiency (IMD) pathways, which are the front line of defence against infection by microorganisms. No compound showed a remarkable effect on the Toll pathway, while 21 inhibited the IMD pathway with an IC50 of 0.27 ± 0.01 μM (Fig. 6a). This result implies that the γ-pyrene moiety that is observed in 21 might have a role in the activity and that the THF ring at the C5 unit enhances the effect, since 13 tend to show a moderate inhibitory effect on the IMD pathway. Then, we evaluated the insecticidal activity of the most potent cytotoxic 5 and IMD pathway inhibitor 21 by using adult Drosophila (Fig. 6b). The result clearly indicates that 5 efficiently paralysed adult Drosophila 1 h after injection, implying that subglutinol A (5) is one of the virulence factors of entomopathogenic Metarhizium fungi against insects. Thus, we successfully found unique biological activity against Drosophila in 5 and 21, which may potentially be developed as insecticides.

**Anti-HIV assay.** We examined the inhibitory activities of DDPs against various species of bacteria (both gram-negative and gram-positive bacteria), fungi and viruses. No DDPs showed significant anti-bacterial or anti-fungal activities. In contrast, preliminary
Screening of amyloid Aβ 42 aggregation inhibitors. The aggregation of the 42-mer-amyloid β (Aβ42) is involved in the pathogenesis of Alzheimer’s disease (AD)\(^\text{63}\). A nucleation-dependent polymerization model that is composed of nucleation and elongation phases is generally accepted as the aggregation mechanism of Aβ42\(^\text{64}\). In the nucleation phase, the monomer of Aβ42 gradually forms low-molecular-weight oligomers (called nuclei), which cause synaptotoxicity and memory loss\(^\text{65}\).

We evaluated the ability of DDPs to inhibit Aβ42 aggregation by using the Th-T (thioflavin-T) assay, which is a conventional and useful method for quantifying Aβ aggregates (Supplementary Fig. 47). Consideration of the preliminary results and SARs, we picked up 12, 13, 14, 20, and 21, and tested them again (Fig. 7a). The fluorescence of Aβ42 and the Th-T complex began to increase after 4 h of incubation, and only 13 and 21 delayed the nucleation phase of Aβ42 until 6 and 13 h, respectively. In contrast, the congener of 13 and 21 regarding the primary alcohol or methyl ester group, such as 12, 14, and 20 showed no delay even though 12, 14 and 20 partly suppressed the elongation of Aβ42 in TEM analysis. Formation of typical amyloid fibrils of Aβ42 was strongly prevented only by 13 and 21, leading to the fragmentation of fibrils (Fig. 7b).

To further analyse the inhibitory mechanism, we subjected a mixture of Aβ42 and 21 after 1 h of incubation at room temperature to liquid chromatography/quadrupole time-of-flight mass spectrometry (LC/Q ToF-MS). The possible peaks of Aβ42-21 adducts were detected, and then the mass envelopes at +7, +6, +5, and +4 charge distribution corresponded to the Michael adduct based on the loss of 32 Da (Supplementary Figs. 48b–d, 49). However, no such adduct was found in the presence of 20 as a negative control (Supplementary Fig. 48a). To identify the specific amino acids involved in the interaction of 21 with the Aβ monomer, we performed LC–MS/MS analysis with collision-induced dissociation (CID) using E22P, M35ox-Aβ9-35 as a toxic conformer surrogate\(^\text{66,67}\). Based on a large number of fragmented b ions and y ions, Lys16 in Aβ42 (DAEFRHDSGYEVHHQKLVF-FADVSNGKAIIGLMVGVGVI) seemed to be one of the target amino acid residues of 21 (Supplementary Fig. 50).

To investigate the reaction mechanism between the highly oxidized γ-pyrene observed in 13 and 21, and the lysine residue in Aβ42 with a decrease of 32 Da, we carried out the following experiment. Benzyl-protected lysine 28 was reacted with 13 and 21 as well as the structurally related 14 and 20 in THF at room temperature (Supplementary Note 30), and each reaction was monitored by HPLC analysis at 1 h and 10 h after mixing (Supplementary Fig. 51a). In the reaction of 13, two new peaks gradually appeared at 1 h later, and 13 completely converted to the new peaks 26 and 27 (Supplementary Fig. 51b) in 10 h. LCMS analysis showed the same molecular ion peak of 26 and 27 at m/z 863 [M + Na]\(^+\), which suggested that 13 connected to 28 with a decreasing of 32 Da in the same situation described above. The reaction of 21 proceeded in the same manner as that of 13, whereas 14 and 20 did not afford any adducts under the same reaction conditions (Supplementary Fig. 51), suggesting that the γ-pyrene structure, including a hydroxyl group and carboxyl methyl ester, was necessary for the reaction. After scaling up the reaction, we isolated 26 and 27, which are in the C-21 epimer relationship, and determined their structures (Fig. 7c, Supplementary Note 31). From their structures, we proposed the following reaction mechanism: 13 took the ortho ester form, which increased electrophilicity at C-24, and the amine in lysine performed a nucleophilic attack at C-24 with elimination of the methoxy group followed by electron transfer to form 26 and 27 (Fig. 7c and Supplementary Fig. 52). However, the pyrene moiety

---

**Fig. 6 Biological activity against Drosophila.** a Inhibitory effect of compound 21 on the IMD pathway Drosophila (2)mbr cells were stimulated with heat-killed E. coli and the activity wells of the Attacin promoter, a read-out of the activation of the IMD pathway, was monitored by a luciferase assay. Data are means ± SEM of triplicate wells from a single experiment and are representative of two independent experiments. b Toxic effect of compound 5 on Drosophila adult. Adult flies (ca. 1 mg each) were injected with the indicated compounds (3.5 ng each), and the flies that did not show movement or only faint shaking of their legs were counted as unmoving flies. Data are the means ± SEM of triplicate samples from a single experiment and are representative of two independent experiments. Source data underlying Fig. 6a, b are provided in a Source Data file.
did not react with the primary thiol in the cysteine derivative, the hydroxyl group in the serine derivative or the guanidine group in the arginine derivative (Supplementary Fig. 53 and Supplementary Note 32). Thus, we successfully discovered that DDPs markedly inhibited Aβ aggregation in the nucleation phase via a lysine-selective binding motif.

**Discussion**

Herein, we report the advantage of the synthetic biology approach based on heterologous biosynthesis coupled with genome mining for rationally expanding the chemical space of biologically active natural products. In this study, we focused on fungal DDPs as a natural product family including a privileged structure and aimed to produce a diverse set of DDPs. Genome mining revealed putative gene clusters for DDPs distributed in five fungal genera and bioinformatics analyses were performed to draw the five treelike DDP biosynthetic pathways. Stepwise reconstruction of all the pathways in *A. oryzae* allowed us to make transformants corresponding to all the intermediates and end products in the pathways, resulting in the isolation of all the DDPs produced in the native DDP pathways. From their structures, we determined the function of all the modification enzymes in the pathways. Subsequently, we designed four extended pathways by using the five native pathways as tool for combinatorial biosynthesis. This combinatorial biosynthesis enabled us to access non-native analogues that equipped further modifications than those of compounds produced via native pathways. Finally, we achieved heterologous biosynthesis of 22 DDPs, including 15 new analogues, which were included intermediates, end products, shunt products and additionally modified analogues in *A. oryzae*.

We then screened the DDP-focused library with various biological activity assays. DDPs produced in this study shared same skeleton and showed similar antiproliferative activities against cancer cells and inhibition of mitochondrial complex III each other. However, interestingly, small structure differences in each DDP gave unique functions of markedly reducing CSC-like populations in MCF-7 bulk cells (11, 12 and 22), potent inhibitory activity of an insect innate immune system, the IMD pathway (21), potent paralysing activity against adult *Drosophila* (5), selectively inhibited HIV proliferation (20), and Aβ aggregation in the nucleation phase through trapping lysine residue on highly modified γ-pyrene motif (13 and 21). Notably, compounds produced via extended pathways indeed showed unique biological activities. Thus, this study showcases the capability of combinatorial synthetic biology in acceleration of drug discovery.

In the post-genomic era, a synthetic biology approach is undoubtedly one of the most powerful methods to achieve not only the generation of natural products from gene resources but also the rational expansion of bioactive natural product chemical space. The gene resources available in the method are rapidly increasing; therefore, the method would infinitely expand the chemical diversity of natural products and their analogues. In addition, the method solves the supply issues and permits natural products to be subjected to enough biological evaluations. When the method becomes advanced and widely used, natural products will be easier available for drug discovery and definitely increase the opportunity to develop natural products as drug seeds.

**Methods**

**General methods.** Polymerase chain reaction (PCR) was performed using a TaKaRa PCR Thermal Cycler Dice® Gradient (TaKaRa Bio) and Thermal Cyclers LifeECO (Nippon Genetics). Oligonucleotide primers for PCR were purchased from Hokkaido System Science Co., Ltd. (Hokkaido, Japan) and listed in Supplementary Table 4. Analytical and preparative TLC was performed on silica gel 60 F254 (Merck) and RP-18 F254S (Merck). Column chromatography and flash chromatography were carried out on silica gel 60 N (100–210 μm, Kanto Chemical) and silica gel 60 N (40–50 μm, Kanto Chemical), respectively. NMR spectra were

---

**Fig. 7 Inhibitory activity of Aβ aggregation.** a Selected results of the Th-T assay (n = 3). b TEM analysis of typical amyloid fibrils formed by Aβ42. Scale bar = 100 nm. Representative micrographs are selected from at least six micrographs taken in one grid. c Hypothetical reaction mechanism of 13 and lysine derivative 28 (13 was reacted with 28 (5.0 eq) in THF at r.t. for 10 h to give adducts 26 and 27). Source data underlying Fig. 7a, b are provided in a Source Data file.
recorded on a Burker AVANCE III spectrometer. Chemical shifts for 1H and 13C NMR are given in parts per million (δ) relative to tetramethylsilane (δH 0.00) and residual methyl groups (δC 77.00) for CDCl3, as internal standard. Reverse phase HPLC analysis was performed on HITACH LaChrom Elite series equipped with a L-2130 pump, L-2200 autosampler, L-2455 Diode Array Detector, and D-2000 system manager. LC-MS analysis was performed on HITACHI Chromaster series equipped with 5110 pump, 5430 Diode Array Detector, 5610 MS Detector, MSD system manager. High performance liquid chromatography was performed on JASCO V-730 spectrophotometer. IR spectra were recorded on JASCO-FT/IR-4200 spectrometer. Optical rotations were recorded on JASCO-P-1030.

**Genome mining and bio-informatics analyses of DDP pathways.** Draft genome sequence of *A. sacchari* Kumo-3 was obtained in previous study (Supplementary Data 1)99. Draft genome sequences of *F. graminearum* PH-1, *M. phusoides* MS6, *C. gigaschissus* IM1349063 and *M. anisopliae* E6 were obtained from National Center for Biotechnology Information (NCBI). Genome mining was performed by Protein BLAST search against *A. sacchari* Kumo-3 genome sequence and the NCBI non-redundant database and 2ndFind program (http://biosys.nh.uh.ipp/2nFind/).

The results of bioinformatics analyses are shown in Supplementary Tables 1-3.

**Fungal strains used as genomic DNA donor.** *Arthrinium sacchari* (strain Kumo-3) was isolated from a spider previously99. Genomic DNA of *Fusarium graminearum* 50218 was obtained from the Medical Mycology Research Center, Chiba University (Chiba, Japan). *Macrophomina phaseolina* 73717M and *Metharziales anisopliae* NBRCC 13233 were obtained from the Biological Resource Center, National Institute of Technology and Evaluation (Chiba, Japan). *Colletotrichum higginsianum* MAFF 305635 was obtained from the Genetic Resource Center, National Agriculture and Food Research Organization (Ibaraki, Japan).

**Heterologous host.** *Escherichia coli* DH5α (Competent Quick, TOYOBO) was used for cloning experiments. *Aspergillus oryzae* NRAS1, a quadraple auxotrophic mutant (niaD<sup>−</sup>, Δac, ΔargH, adeA<sup>−</sup>) was used as the host for fungal expression.

**Preparation of fungal genomic DNA.** *A. sacchari* Kumo-3, *M. phaseolina* NBRCC 7317, *M. anisopliae* NBRCC 103233, and *C. gigaschissus* MAFF 305635 were cultivated on Potato dextrose agar (PDA) (2.4% Potato dextrose broth (Difco), 1.5% agar). The spores and mycelium from the plate were inoculated into 60 mL of Potato dextrose medium (2.4% Potato dextrose broth). After several days cultivation at 30 °C (reciprocal shaking), the mycelia of each strain were collected by filtration, washed with water, and frozen at –80 °C. The frozen mycelium was ground to fine powder, suspended in TE buffer (pH 8.0), and equal volume of lytic buffer (2% SDS, 0.1 M NaCl, 10 mM EDTA, 50 mM Tris-HCl was added. After incubation at room temperature for 5 min, the supernatant was extracted with phenol: chloroform: isoamyl alcohol (25:24:1) solution (pH 7.9). After ethanol incubation at room temperature for 5 min, the supernatant was extracted with 1.5% agar). The spores and mycelium from the plate were inoculated into 60 mL of *M. anisopliae* (strain Kumo-3), and the suspension was combined with molten soft agar (3.5% Czapek-Dox broth (Difco), 0.5% Casamino acids vitamin assay medium (3.5% Czapek-Dox broth, 0.8 M NaCl, 0.5% maltose monohydrate (Nacalai Tesque), 0.01% adenine (for adenine auxotrophic strains)). Transformation of *A. oryzae* transformant. We obtained about four to ten transformants per one-transformation experiments. Thus, we checked gene insertion of all the transformant by PCR. Target gene integration into the *A. oryzae* transformants were confirmed by PCR with the genomic DNA template and primers in Supplementary Table 4. Genomic DNA of the transformant was extracted as follows; small pieces of mycelium of the *A. oryzae* strain were collected from an agar plate after several days growth at 30 °C and incubated at 95 °C in TE (pH 8.0) buffer for 10 min. The genomic DNA solution was diluted with water to prepare the PCR template at appropriate concentration.

**Cultivation and metabolite analysis of *A. oryzae* transformants.** All the transformant that pass the gene integration check were cultivated and analysed their metabolite profiles by reverse phase HPLC and LC-MS analyses. Each transformant was inoculated into 60 mL of CPS medium (3.5% Czapek-Dox broth, 0.3% casein peptone (Nacalai Tesque), 0.3% meat peptone (Nacalai Tesque), 0.3% soy peptone (Nacalai Tesque), 0.01% trilaurin triolein, 2% soluble starch (Nacalai Tesque), 1% maltose monohydrate (Nacalai Tesque), 0.01% adenine (for adenine auxotrophic strains)) or 1/2 CPS medium (1.75% Czapek-Dox broth, 0.15% casein peptone (Nacalai Tesque), 0.15% meat peptone (Nacalai Tesque), 0.15% soy peptone (Nacalai Tesque), 0.1% soluble starch (Nacalai Tesque), 0.15% sodium acetate (Nacalai Tesque), 0.01% adenine (for adenine auxotrophic strains)). About 1 week (6–10 days) after cultivation at 30 °C, 150 rpm, the culture media and mycelium were separated by filtration. The culture media (3 mL) was extracted with EtOAc (2 mL), and the organic layer was concentrated in vacuo. The residue was then dissolved in 150 μL of MeOH to prepare sample for reverse phase HPLC and LC-MS analysis. The freeze-dried mycelium (20 mg) was extracted with MeOH (1 mL), and the extract was concentrated in vacuo. The residue was then dissolved in 100 μL of MeOH to prepare sample for HPLC and LC-MS analysis. For the HPLC and LC-MS analysis, 10 μL of the samples were used. The HPLC analysis was performed on COSMOSIL 5C18 Packed Column (4.6 mm I.D. × 150 mm, 5 μm). The mobile phase consisted of 0.01% triethylammonium acetate, 1% acetonitrile containing 0.1% formic acid and water containing 0.1% formic acid (0:20:80:1.5 min; 20:80:1.5:11.5 min; a linear gradient from 20:80 to 100:0, 11.5–20 min: 100%) at a flow rate of 1.0 mL/min<sup>−</sup>. The LC-MS analysis was performed on COSMOSIL 5C18 Packed Column (4.6 mm I.D. × 150 mm, Nacalai Tesque) with acetonitrile containing 0.1% formic acid and water containing 0.1% formic acid (0:2 min: 80:20, 2–12 min: a linear gradient from 20:80 to 100:0, 12–20 min: 100%) at a flow rate of 1.0 mL/min<sup>−</sup> using positive mode electrospray ionization.
for 1 week at 30 °C, and the culture media and mycelium were separated by filtration. The culture media was extracted with EtOAc two times, and the organic layer was concentrated and concentrated in vacuo to give crude extract. The freeze-dried mycelium was extracted with EtOAc (20% MeOH), and the extract was concentrated in vacuo. The mycelia extract was then dissolved in MeOH (0.1% water), washed with hexane three times, concentrated in vacuo to give crude extract. Each of the crude extract of culture and mycelium was fractionated by silica gel column chromatography, and subjected to further purification by flash chromato- graphy and preparative TLC (SiO₂ or ODS). All the DDP structures were fully determined by spectral analyses (Supplementary Figs. 8–10, 14, 15, 18, 19, 22, 23, 26, 27, 32, 33, 37, 38, 41, 42, 52, 54–77, Supplementary Tables 7–18, 25 and Supplementary Notes 3–6, 8, 10, 12, 14, 18, 20, 22, 31).

Evaluation of biological activities of DDP-focused library. Purified DDPs was dissolved in DMSO (Specially Prepared Reagent, Nucleose and Protease tested, NACALAI, 09659-14) to prepare 5 mM solutions. The Each DDP solution (5 mM) was used for antiproliferative effects on cancer cell lines (Supplementary Notes 23, 24, 25, 26), biological activity against Drosophila (Supplementary Note 27), anti-HIV assay (Supplementary Note 28) and screening of amyloid Ab42 aggregation inhibitors (Supplementary Notes 29).

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability
The data supporting the findings of this work are available within the paper and its Supplementary Information files. A reporting summary for this Article is available as a Supplementary Information file. The data sets generated and analyzed during this study are available from the corresponding author upon request. The sequences data of each DDP is deposited at DDBJ/EMBL/GenBank. The data supporting the findings of this work are available within the paper and its Supplementary Information files. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

References
1. Koehn, F. E. & Carter, G. T. The evolving role of natural products in drug discovery. Nat. Rev. Drug Discov. 4, 206–220 (2005).
2. Newman, D. J. & Craig, G. M. Natural products as source of new drugs from microorganisms. Nat. Prod. Rep. 33, 26–53 (2016).
3. Matsuda, Y. & Abe, I. Biosynthesis of fungal meroterpenoids. Nat. Prod. Rep. 33, 331–369 (2016).
4. Masschelein, J., Jenner, M. & Challis, G. L. Antibiotics from Gram-negative bacteria: a comprehensive overview and selected biosynthetic highlights. Nat. Prod. Rep. 34, 712–783 (2017).
5. Kellenberger, E., Hofmann, A. & Quinn, R. J. Similar interactions of natural products with biosynthetic enzymes and therapeutic targets could explain why nature produces such a large proportion of existing drugs. Nat. Prod. Rep. 28, 1483–1492 (2011).
6. Cosentino, G., Grattan, M., Muresan, S., Backlund, A. & Oprea, T. L. Novel chemical space exploration via natural products. J. Med. Chem. 52, 1953–1962 (2009).
7. Clardy, J. & Walsh, C. Lessons from natural molecules. Nature 432, 829–837 (2004).
8. Over, B. et al. Natural-product-derived fragments for fragment-based ligand discovery. Nat. Chem. 5, 21–28 (2013).
9. Maier, M. E. Design and synthesis of analogues of natural products. Org. Biomet. Chem. 13, 5302–5343 (2015).
10. Robles, O. & Romo, D. Chemo- and site-selective derivatizations of natural products enabling biological studies. Proc. Natl Acad. Sci. USA 111, 12354–12359 (2014).
11. Awakawa, T. et al. Reprogramming of the antimycin NRPS-PKS assembly lines inspired by gene evolution. Nat. Prod. Rep. 33, 351–371 (2016).
12. Jørgensen, L. et al. 14-step synthesis of (−)‐α-pyrone and (+)‐β-pyrone: discovery of fungal natural products. ACS Chem. Biol. 8, 257–267 (2013).
13. Kikuchi, H. et al. New diterpene pyrone-type compounds, metarhizin A and B, isolated from entomopathogenic fungus, Metarhizium flavoviride and their inhibitory effects on cellular proliferation. Tetrahedron 65, 469 (2009).
14. Lee, J. C., Lobkovsky, E., Plaim, N. B., Strobel, G. & Clardy, J. Subglutinols A and B: immunosuppressive compounds from the endophytic fungus Fusarium subglutinans. J. Org. Chem. 70, 7076–7077 (2005).
15. Lee, W. G. et al. Immunosuppressive effects of subglutinol derivatives. J. Antibiot. 58, 397–404 (2005).
16. Lai, K. et al. Integrated compound profiling screens identify the mitochondrial electron transport chain as the molecular target of the natural products manzamine, sesqiquin, and atrazine. ACS Chem. Biol. 8, 257–267 (2013).
17. Gohbara, M. et al. Structures of collettotrichin and colletrotonin B, phytotoxic metabolites from Colletotrichum nicotianae. Agric. Biol. Chem. 42, 1037–1043 (1978).
18. Zhan, F. & Danishefsky, S. J. An efficient stereoselective total synthesis of (−)‐sesquicillin, a glucoconorcinoglut. Angew. Chem. Int. Ed. 41, 1344–1347 (2002).
19. Kim, H. et al. Stereoselective synthesis and osteogenic activity of subglutinols A and B. J. Am. Chem. Soc. 131, 3192–3194 (2009).
20. Oguchi, T., Watanabe, K., Ohkubo, K., Abe, H. & Katou, T. Enantioslective total synthesis of (−)‐subglutinols A and B: potential immunosuppressive compounds. Chem. Eur. J. 15, 2826–2845 (2009).
21. Merchant, R. et al. Divergent synthesis of pyrrole diterpenes via radical cross coupling. J. Am. Chem. Soc. 140, 7462–7465 (2018).
Acknowledgements

The authors thank Prof. K. Gomi (Tohoku University) and Prof. K. Kitamoto (The University of Tokyo) for providing the expression vectors and the fungal strain. The authors thank Prof. T. Kuroda (Hiroshima University) and Prof. D. Hagwara (Tusukuba University) for preliminary screening of anti-bacterial and anti-fungal compounds. The JFCR39 cancer cell line panel assays were kindly performed by Molecular Profiling Committee, Grant-in-Aid for Scientific Research on Innovative Areas "Platform of Advanced Animal Model Support (AdAMS)" from The Ministry of Education, Culture, Sports, Science and Technology, Japan (KAKENHI 16H06276 to S.D.). This work was supported by JSPS KAKENHI (grant no. 17H05055, 17H05428, 18K19388 and 19H04624 to T.A., 16H06194 to K.M. and 17H04107 to Y.K.) from Japan Society for the Promotion of Science (JSPS), the Naito Foundation and Astellas Foundation for Research on Metabolic Disorders to T.A.

Author contributions

T.A. conceived this study, and designed the experiments. T.A. and K.T. analysed all the DDP clusters and performed all the synthetic biology experiments and characterized all the DDVs. T.A. and K.T. also performed chemical reactions. S.S. and A.K. contributed to construct A. oryzae transformants and analysed their metabolite profiles by HPLC analyses. K.M. and K.I. designed, performed and analysed the assay for amyloid Aβ42 aggregation inhibition. M.M. and H.M. designed, performed and analyzed assay for inhibitory activities on mitochondrial respiratory complex. S.D. designed, performed and analysed the assay for antiproliferative effects across the panel of JFCR39 cancer cell lines. A.H., E.S. and T.K. designed, performed and analysed the assay using Drosophila. K.K., H.H. and E.N.K. designed, performed and analysed the assay for anti-viral activity. N.H. and Y.K. designed, performed and analysed the assay for anti-CSC activity. T.A. wrote this manuscript with input from all authors. All authors have given their approval of the final version of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41467-020-15664-4.

Correspondence and requests for materials should be addressed to T.A.

Peer review information Nature Communications thanks Ren Xiang Tan, and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2020