



Efficient exploration of terpenoid biosynthetic gene clusters in filamentous fungi

Yujie Yuan^{1,4}, Shu Cheng^{1,4}, Guangkai Bian^{ID 1,4}, Pan Yan¹, Zhengning Ma¹, Wen Dai², Rong Chen¹, Shuai Fu¹, Huiwen Huang¹, Haoming Chi¹, Yousheng Cai¹, Zixin Deng^{1,3} and Tiangang Liu^{ID 1,3}

Terpenoids, critical components of human medicine, are the largest family of natural products. Fungi are an important source of terpenoids, but many of the corresponding biosynthetic gene clusters (BGCs) are silent in laboratory conditions. Strategies such as homologous activation and heterologous expression were usually used to active a single cluster, making them low efficiency. Here we developed an automated and high-throughput (auto-HTP) biofoundry workflow using *Aspergillus oryzae* as a chassis that enables efficient genome mining, characterization of BGCs and identification of bioactive fungal terpenoids. We simultaneously refactored 39 BGCs into 208 engineered strains, producing 185 distinct terpenoids. An anti-inflammatory screen returned the sesquiterpenoid mangicol J; re-examination of our engineered strains revealed the likely biosynthetic pathway. Finally, we optimized the mevalonate pathway in *A. oryzae* to provide a more efficient chassis for overproduction of terpenoids. The auto-HTP biofoundry workflow together with the optimized *A. oryzae* chassis can accelerate the discovery and development of terpenoid natural products.

Natural products discovered in fungi represent an indispensable source of pharmaceutical agents^{1,2}. Traditional strategies to mine microbes for molecules have revealed a vast number of bioactive natural products, including the potent antibiotic penicillin, the antihypercholesterolemic agent lovastatin and the immunosuppressant cyclosporine^{3–5}. The research scopes of natural products have expanded from exploring high-abundance molecules that could be fractionated from grown cultures, to the development and use of bioinformatic analysis and computational tools, such as antiSMASH and SMURF^{6,7}, that have helped to reveal massive transcriptionally silent or cryptic biosynthetic gene clusters (BGCs). Strategies to activate these BGCs and initiate production of their bioactive compounds are critical to release the potential power of nature's chemical repertoire⁸. For example, researchers have demonstrated that manipulation of global or pathway-specific regulators^{9,10}, application of elicitors¹¹ and gene editing mediated by CRISPR–Cas9 (refs. ^{12,13}) can be used to activate silent BGCs and unlock the biosynthesis of new natural products. However, application of these strategies can be restricted by difficulties in establishing a genetic manipulation system for a specific microbe, and the difficulties of distinguishing and purifying a target product from unwanted products. Moreover, many efforts thus far have focused on a specific molecule, pathway or organism of interest; however, the extensive unexplored chemical space suggests that new approaches should be versatile and high-throughput.

BGC refactoring provides one such approach, stripping out native regulation of pathways to enable metabolite biosynthesis¹⁴. Heterologous expression is the predominant approach for BGC refactoring, enabling mechanistic studies of the resulting polyketides, peptides and terpenoids^{15,16}. Other emerging strategies, including bioinformatics and computational frameworks, artificial chromosomes, metabolic scoring-based approaches and the HEx (heterologous expression) platform, have further advanced our

ability to mine cryptic BGCs^{17–20}. However, low efficiency and insufficient quantities of compounds for structural characterization remain challenges in the field. Further adoption of metabolic engineering principles, in which efficient chassis are constructed that accumulate relevant precursors, can improve compound yields. However, existing chassis built from *Escherichia coli* and *Saccharomyces cerevisiae*^{20–23} are not well suited to production of fungal natural products, as they poorly express membrane-localized cytochrome P450 enzymes (CYP450s), lack the requisite redox partner proteins and do not contain appropriate machinery to identify and splice the intron-rich genes of filamentous fungi. As such, these chassis have primarily found success in the expression of terpene cyclases, to synthesize core hydrocarbon skeletons^{21,24}.

The filamentous fungi *A. oryzae* (AO) is generally recognized as safe and provides an improved cellular context to explore fungal natural products²⁵. Indeed, it is an ideal chassis for the mining of individual terpene cyclases^{26,27}, or entire BGCs from filamentous fungi^{28–30}, enabled by CRISPR–Cas9-mediated gene editing^{31–33}. Even in AO, however, the limited supplies of precursors (for example, isopentenyl pyrophosphate and dimethylallyl pyrophosphate) make it challenging to acquire enough terpenoids for biochemical and functional studies, especially for late derivatives that require multiple biosynthetic steps.

Herein, we used the AO NSAR1 as a v.1.0 chassis, to develop an automated and high-throughput (auto-HTP) biofoundry workflow as a new genome mining strategy (Fig. 1). We used auto-HTP biofoundry workflow to reconstitute 39 terpenoid BGCs, and evaluated the activity and structural novelty of the terpenoids produced by 208 engineered strains. Functional screening revealed an anti-inflammatory compound, mangicol J, with activity in vitro and in vivo, while our engineered strains provided a facile framework to quickly identify the enzymes involved in synthesis of mangicol J and its family members. Finally, we developed a second-generation

¹Key Laboratory of Combinatorial Biosynthesis and Drug Discovery, Ministry of Education and School of Pharmaceutical Sciences, Wuhan University, Wuhan, China. ²Department of Clinical Laboratory, Renmin Hospital of Wuhan University, Wuhan, China. ³Hubei Engineering Laboratory for Synthetic Microbiology, Wuhan Institute of Biotechnology, Wuhan, China. ⁴These authors contributed equally: Yujie Yuan, Shu Cheng, Guangkai Bian.

e-mail: liutg@whu.edu.cn

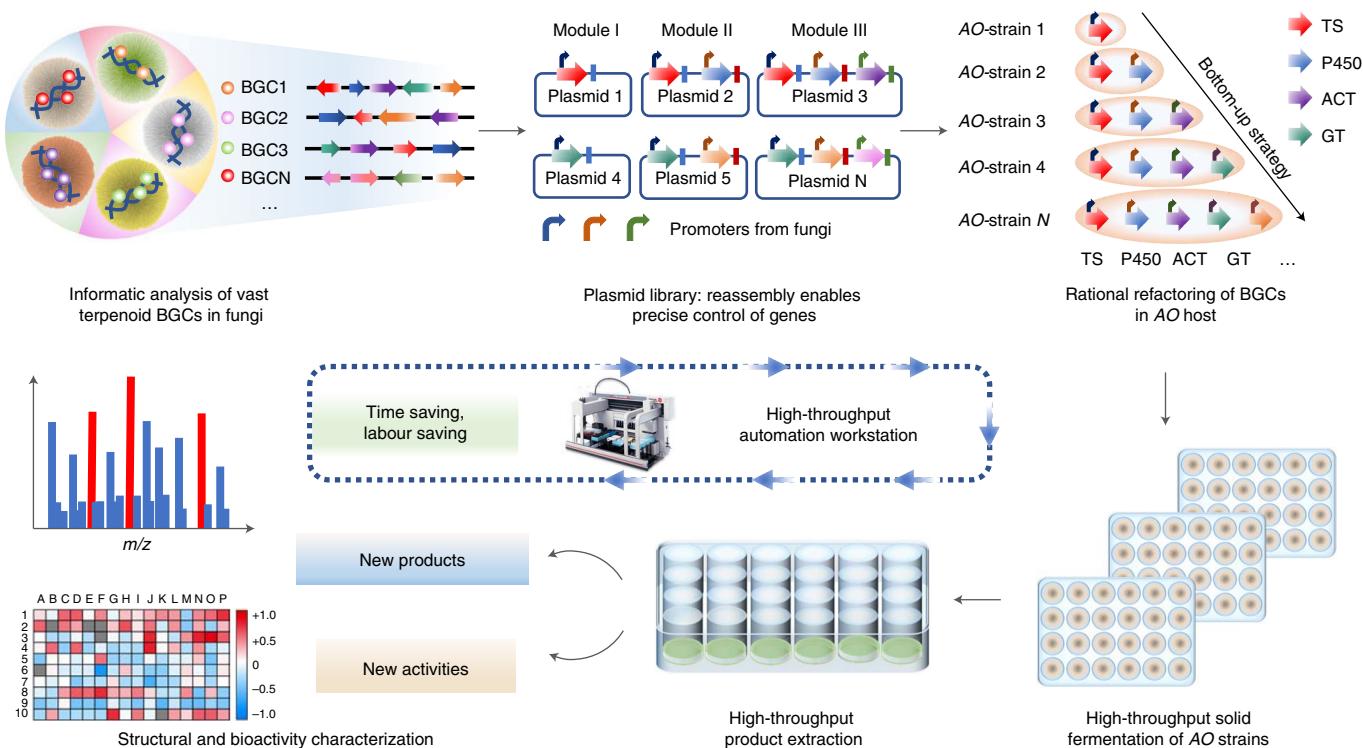


Fig. 1 | Overview of automated and high-throughput (auto-HTP) biofoundry workflow. There included several modules in the workflow, from bioinformatic analysis of fungal terpenoid BGCs, plasmid library construction by automated yeast assembly, rational refactoring of BGCs in *AO* chassis to high-throughput fermentation of *AO*-strains, high-throughput product extraction and structural and bioactivity characterization. TS, terpene synthase; P450, cytochrome P450 enzyme; ACT, acetyltransferase and GT, glycosyltransferase.

AO chassis (v.2.0) with an improved titre of metabolic precursors, which serves as a universal and efficient platform to overproduce terpenoids such as mangicol J. This combined strategy represents an efficient approach for the large-scale genome mining of terpenoids in filamentous fungi, and also provides a general approach for the mining of other natural products.

Results

Bioinformatic analysis and principles for refactoring of terpenoid BGCs. *AO* NSAR1 (*niaD*⁻, *sC*⁻, *ΔargB*, *adeA*⁻), an auxotrophic strain that has been widely used for heterologous expression of genes from filamentous fungi^{25,34}, was selected as the host for mining terpenoid BGCs found in five in-house genome sequenced filamentous fungi, namely *Colletotrichum gloeosporioides* ES026, *Alternaria alternata* TPF6, *Fusarium graminearum* J1-012, *Trichoderma viride* J1-030 and *Aspergillus flavipes*. These five fungi were initially predicted to contain 54 terpenoid BGCs on the basis of antiSMASH analysis. Putative BGCs were ruled out if they contained terpene synthases missing known conserved motifs (the DDXD/E and NSE/DTE motifs for class I and the DXDD motif for class II terpene synthases) or with incomplete functional domains, leaving 39 BGCs predicted to produce sesqui-, di- and sesterterpenoids (Fig. 2a).

Next we constructed a plan to refactor the 39 clusters. However, the heterologous expression of terpenoid BGCs in *AO* is often hampered by challenges in recognizing the native promoters of BGCs, leading to random selection of promoters without consideration of how to precisely control the expression of the functional genes. Here we characterized a series of promoters that could be used in place of the native sequences. We used β-glucuronidase as a readout to determine the strength of these promoters, in the order of *hlyA* > *oliC* > *amyB* > *glaA* > *enoA* > *gpdA* > *agdA* > *trpC* > *alcA* (Extended Data Fig. 1). We next classified the functional genes

in the BGCs into upstream terpene synthase modules, in which cyclization of oligoprenols produces the polycyclic hydrocarbon or alcohol with multiple stereocenters, midstream oxidation modules, in which the core is oxidized to form double bonds, carbonyls and alcohol groups, and downstream modules, in which optional post-functionalization steps (for example, acylation, glycosylation and so on) form the final complex architecture³⁵. Combining these results provided an approach to replace the regulatory genes, placing functional genes under the control of the strong constitutive (*hlyA*) and inducible (*amyB*, *glaA*) promoters. With the help of the auto-HTP biofoundry, genes in the same module were either placed in separate plasmids or coconstructed into a plasmid for all possible combinations. Following the combination of plasmids from different modules, we refactored the BGCs into a strain library, which contained the terpene synthase gene as the basic start point and then individually added the other downstream genes from the same BGC (Extended Data Fig. 2). This approach would enable us to not only decipher the function of each (partial) cluster on the basis of the resulting compounds produced, but also identify the specific role of each new enzyme and examine the bioactivity of the resultant terpenoids.

An automated and high-throughput biofoundry workflow for terpenoid BGCs. Biofoundries are widely used in DNA assembly, enzyme evolution and strain engineering^{36–39}. We envisioned that their application to natural product biosynthesis would enable mining of a large number of cryptic BGCs in an automated and high-throughput way. The auto-HTP biofoundry workflow uses advanced liquid-handling technologies to accelerate the processes of PCR amplification, plasmid construction using yeast homologous recombination and transformation of *AO* protoplasts to refactor the terpenoid BGCs (Fig. 2b and Extended Data Fig. 3). We used

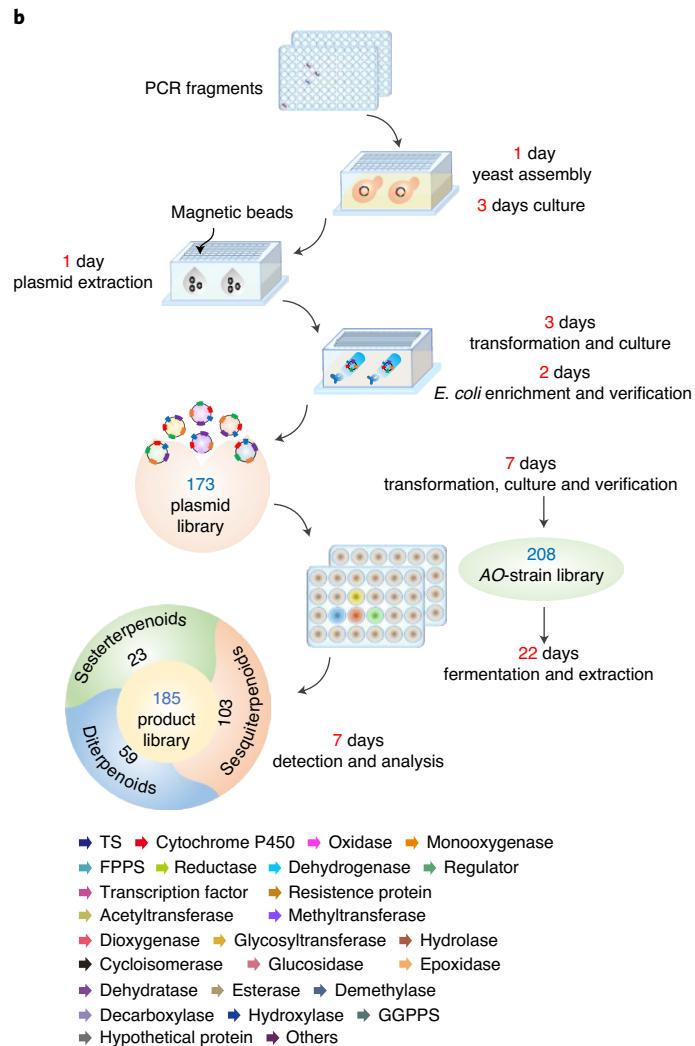
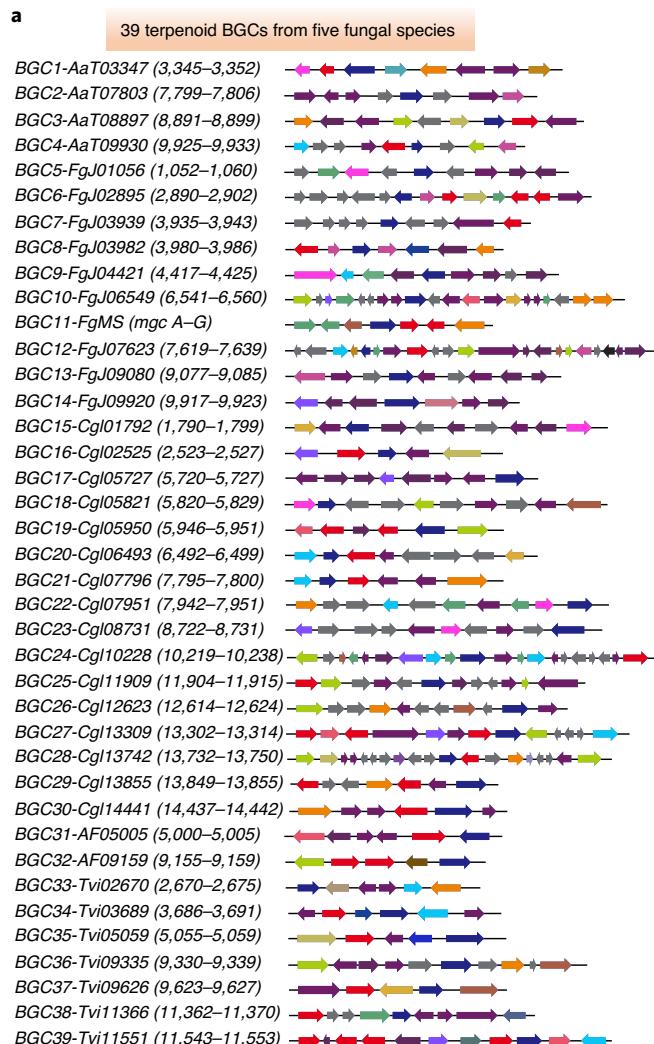


Fig. 2 | Fungal terpenoid BGCs refactory using auto-HTP biofoundry workflow. **a**, The coding genes of 39 BGCs of *AaTxxxx*, *FgJ,xxxx* *Tvixxxx*, *Cglxxxx* and *AFxxxx* were amplified from the genomes of the corresponding fungi *Alternaria alternate*, *F. graminearum* J1-012, *T. viride* J1-030, *C. gloeosporioides* ES026 and *A. flavipes*, respectively. The numbers in brackets (xxxx–xxxx) indicate the starting and ending gene in each BGC. **b**, Refactoring BGCs using an automation workstation.

the Biomek FX[®] Laboratory Automation Workstation (Beckman Coulter) as the automatic platform to carry out PCR amplification, PCR product recovery and plasmid extraction, as well as transformation of *E. coli*, *S. cerevisiae* and AO protoplasts (Supplementary Figs. 1–6). The Molecular Device Qpix 460 was used to pick up colonies. Using the auto-HTP biofoundry workflow, we acquired 589 fragments (98% of the designed sequences) after two rounds of automated PCR amplification; these included the plasmid backbone, promoter, terminator and functional genes from the BGCs (Supplementary Video 1). The remaining nine fragments, probably restricted by long gene fragments and the low specificity of primers, were manually amplified. The fragments were assembled in *S. cerevisiae* via homologous recombination into 173 plasmids (Supplementary Fig. 2 and Supplementary Video 1). The assembled plasmids were transformed into *E. coli* (Supplementary Video 2) for replication and then automatically extracted (Supplementary Video 2). Next, we combined different plasmids into the AO protoplasts, merging the genes in different groups and orders, yielding 208 distinct strains (Extended Data Fig. 2, Supplementary Fig. 5 and Supplementary Video 3). Thus, with the help of the automated biofoundry workflow, we refactored 39 BGCs into sufficient combinations to fully display the metabolic intermediates and final products

of each BGC at one time. Among them, 166 (96%) plasmids and 195 (96%) AO-strains were acquired after two rounds of construction. The remaining seven plasmids and 13 strains, containing multiple genes and large plasmids, respectively, were constructed manually. The entire process from PCR amplification to the final construction of the AO-strains takes 17 days when each step goes smoothly. Finally, the engineered AO-strains were incubated in 24 deep-well plates containing solid rice medium at 30 °C for 2 weeks. Afterwards, compounds were extracted twice from the cultures using acetone and ethyl acetate and concentrated to generate crude extracts. Terpenoid production in 208 AO-strains was detected using gas chromatography-mass spectrometry (GC-MS) and high-resolution-electrospray ionization-mass spectrometry (HR-ESI-MS). A total of 185 distinct terpenoids (**1–185**) from 26 of the 39 BGCs were detected, with 47 only detected by GC-MS, 84 only detected by HR-ESI-MS and 54 detected by GC-MS and HR-ESI-MS simultaneously (Extended Data Fig. 4). Furthermore, 103, 59 and 23 were sesqui-, di- and sesquiterpenoids, respectively (Table 1). Comparison with mass spectrometry analysis of previously purified compounds^{21,23,24,40–44} and the National Institute of Standards and Technology library (<https://webbook.nist.gov/chemistry/>)⁴⁵ indicated that 22 of the terpenoids are previously reported compounds (Extended Data Fig. 5 and

Table 1 | Result summary of BGC reconstruction and detection in this study

Fungal species								
A. alternate TPF6	F. graminearum J1-012			C. gloeosporioides ESO26			T. viride J1-030	A. flavipes
Result summary of each procedure								
Terpenoid BGCs		Plasmids constructed		AO-strains constructed		Products detected		
Total	39	Procedure	Positive rate ^a	Procedure	Positive rate ^a	Sesquiterpenoid	103	
Productive	26	First round	70%	First round	56%	Diterpenoid	59	
Positive rate: 66.7%		Second round	26%	Second round	38%	Sesterterpenoid	23	
PCR fragments (total: 589)		Manual work	4%	Manual work	6%	Total number of products: 185 (34 present with structures)		
Procedure	Positive rate ^a	Total number of plasmids and AO-strains constructed: 173 and 208						
Automation	98%	Time for one round of construction: 17 days						

^aThe positive rate was calculated of the number of positive DNA fragments, plasmids or AO-strains of each round divided by the total number of DNA fragments, plasmids or AO-strains.

Supplementary Table 8). Strains with negative results were rechecked by PCR amplification and refermentation to eliminate any influence caused by the experimental process.

To get a better sense of the remaining chemical diversity produced in our auto-HTP biofoundry workflow, we characterized a representative subset of compounds from different BGC strains by nuclear magnetic resonance (NMR), which was collected and analysed by a Bruker AV III and MestReNova v.5.3.1, respectively. Compounds **6**, **35**, **136** and **144**, produced by strains containing genes from BGC3, BGC4, BGC25 and BGC27, were fully assigned and named as 3,11-eudesmadien-2-one, traversiadiene, (+)-germacrenol and (−)-aristolochene, respectively (Supplementary Figs. 15–20 and Supplementary Tables 6 and 7). Traversiadiene (**35**) is the precursor of the strongly molluscicidal diterpenoid traversianal^{41,42}. This indicates the terpene cyclase from the AaT09930 BGC (BGC4) is traversiadiene cyclase; initial analysis of the refactored AaT09930 strains (Supplementary Fig. 17) provides a solid foundation for further characterization of this gene cluster.

The terpene cyclase of BGC6, FgJ02895, was previously characterized as guaia-6,10(14)-diene cyclase and used in the semi-synthesis of the chemotherapeutic agonist (−)-englerin A³⁵. To clarify whether BGC6 (Fig. 3a) actually produces the plant-derived (−)-englerin A, we examined the terpenoids produced by the refactored strains AO-Y23 to AO-Y33 (Figs. 3b,c and Supplementary Figs. 11–14). Compound **17** (guaia-6,10(14)-diene) was detected from strain AO-Y23, harbouring FgJ02895. Compound **20** was detected from AO-Y24, harbouring FgJ02895/2897, and was characterized by NMR as a new sesquiterpenoid named guaiaoxirane (Fig. 3c and Supplementary Tables 4 and 5). Comparing the structures of compounds **17** and **20** suggests that FgJ02897 is a cytochrome P450 that can catalyse the oxidation of C10 and C14. Compound **28** was detected from strain AO-Y29, harbouring FgJ02895/2897/2898 and characterized by NMR as the C14-acetylated hydroxylated guaiadiol (Fig. 3c and Supplementary Tables 4 and 5). Compound **34** was predicted by liquid chromatography–tandem mass spectrometry to be C14-hydroxy esterified guaiadiol acetate (Supplementary Fig. 14a). These data indicate that, despite starting from the same guaiane-type core-skeleton, the BGC6 and (−)-englerin A pathways are distinct (Fig. 3d and Supplementary Fig. 14b).

Compounds 68–73 were characterized as described below, and the remaining 151 uncharacterized terpenoids were predicted to be methylated, acetylated and oxidized terpenoids with diverse structures based on GC–MS and HR–ESI–MS data, which were collected and analysed by Qual Browser Thermo Xcalibur (Supplementary Figs. 8–50). We viewed this small but varied terpenoid library as a compelling starting point to screen for compound bioactivity. The sequences of the characterized enzymes were deposited in

the National Microbiology Data Center (<https://nmdc.cn/en>) and accession numbers are listed in Supplementary Table 17.

Anti-inflammatory activity of sesterterpenoid mangicol J. We decided to search for molecules that inhibit inflammation, a complex response that can occur when organisms protect themselves from infection. The SARS-CoV-2 pandemic has provided renewed motivation to identify molecules that can provide insights into or ameliorate inflammation. Specifically, infection of COVID-19 has been shown to increase interleukin-6 (IL-6), tumour necrosis factor α (TNF- α) and other inflammatory factors in serum, resulting in the sometimes fatal systemic cytokine release syndrome⁴³. There is some evidence that terpenoids may be privileged scaffolds to engage with and inhibit inflammation⁴⁷, providing additional rationale for our experiments.

We designed a screen in which RAW 264.7 murine macrophage cells would be treated with lipopolysaccharide (LPS) to initiate inflammatory pathways and the release of nitric oxide (NO) monitored as a readout using the Nitric Oxide Assay Kit (S0021, Beyotime); the ability of the crude extracts to decrease levels of NO is then a measure of their anti-inflammatory activity. We first confirmed that the extracts (500 μ g ml^{−1} in dimethylsulfoxide (DMSO)) were minimally or not cytotoxic (Extended Data Fig. 6a). Then we used the auto-HTP biofoundry workflow to test the activity of the extracts (Supplementary Fig. 6 and Supplementary Video 3). These in vitro data indicated that 16% of the fermentation extracts displayed substantial anti-inflammatory activity, reducing NO levels to less than 20% of the untreated control (Fig. 4a), with strains AO-Y52 and Y56–58 demonstrating particularly high activity (Supplementary Table 1). These strains included genes from the FgMS BGC, which have been predicted to produce previously reported anti-inflammatory sesterterpenoid mangicols⁴⁸. We returned to the AO-strains that include genes from this cluster, AO-Y51 to AO-Y62, and purified compounds **68–73** and characterized them by NMR spectroscopy. Compound **71** was identified as the known molecule mangicol E; compounds **68–70**, **72** and **73** were named as the new sesterterpenoids mangicol H–L, respectively (Supplementary Table 2 and 3). In vitro data showed that mangicol J had the highest NO inhibition activity, with an IC₅₀ (the half maximal inhibitory concentration) of $8.29 \pm 0.13 \mu$ M, which is more potent than that of the positive controls N^G-monomethyl-L-arginine (L-NMMA) ($28.48 \pm 6.78 \mu$ M) and indomethacin (IMC) ($40.08 \pm 5.32 \mu$ M) (Fig. 4b). We also tested the purified mangicols up to 100 μ M and confirmed they were not generally cytotoxic (Fig. 4d and Extended Data Fig. 6b).

With these promising results in hand, we evaluated the anti-inflammatory activity of mangicol J in vivo, first using a mouse ear oedema model. We applied mangicol J, IMC or vehicle to the ear, and then applied phorbol myristate acetate (PMA; 2 μ g per ear in

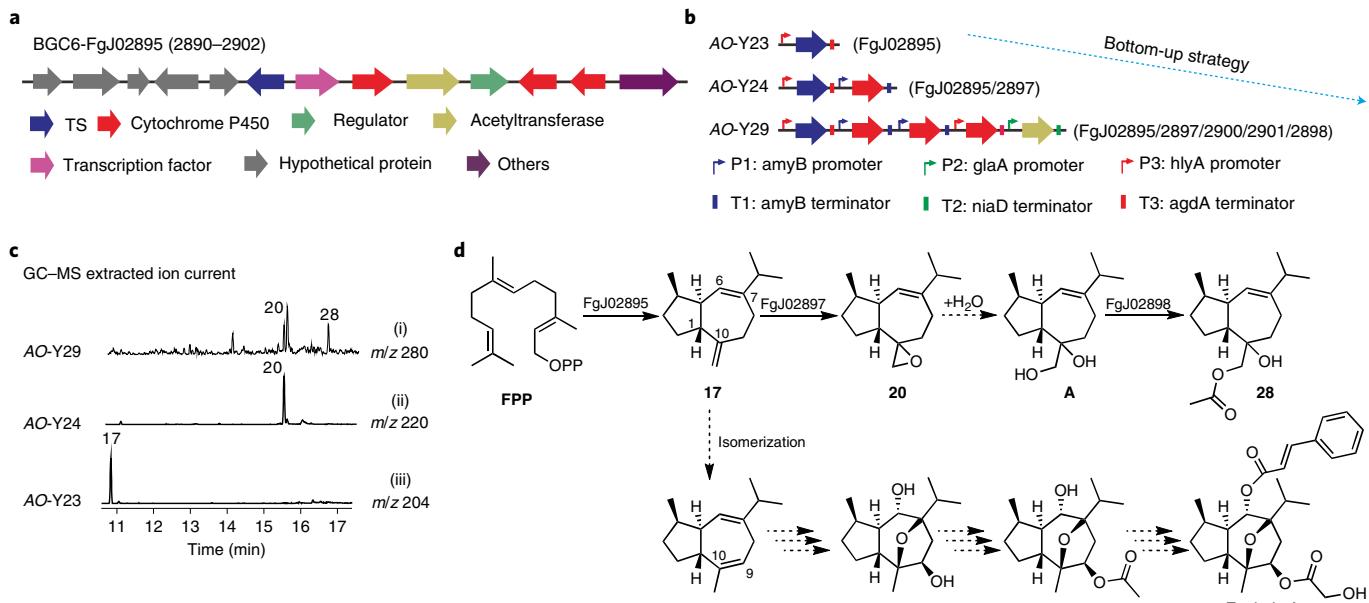


Fig. 3 | Mechanistic characterization of sesquiterpenoids from BGC6-FgJ02895. **a**, Bioinformatic analysis of the FgJ02895 gene cluster. **b**, Reconstitution of the FgJ02895 gene cluster in AO. **c**, GC-MS analysis of molecules produced by AO-strains expressing the FgJ02895 gene cluster. **d**, The proposed biosynthetic pathway for sesquiterpenoids produced by the FgJ02895 gene cluster. FPP, farnesyl pyrophosphate.

20 µl of acetone) to stimulate inflammation. Our data revealed that mangicol J inhibited PMA-induced oedema at the same level as the pharmaceutical agent IMC (58 versus 63%, respectively) (Fig. 4c). We also examined the effect of mangicol J using an LPS-stimulated C57BL/6J mouse model. We observed that 10 mpk of mangicol J significantly decreased the serum levels of IL-6, IL-10, IFN- γ , IL-17A and TNF- α (Fig. 4e–i). Western blot data also showed that mangicol J can decrease the expression level of phosphorylated-signal transducer and activator of transcription 3, which is downstream of IL-6, in human umbilical vein endothelial cells (Supplementary Fig. 7). The anti-inflammatory mechanism of mangicol J in mice and humans requires systematic investigation.

Elucidating the biosynthetic pathway for mangicol J. One strength of the auto-HTP biofoundry workflow is that the refactored gene clusters enable rapid assessment of the specific genes needed to make a compound of interest. As mentioned above, the strains that led us to mangicol J all included genes from the FgMS cluster (Fig. 5a,b and Supplementary Figs. 8–10). Our previous work demonstrated that MgcD (FgMS) is a sesterterpene cyclase, which can catalyse the formation of **64** (mangicdiene)²⁴. The identification of mangicols H–J in AO-Y52 (containing MgcDE) suggests that MgcE is a multifunctional P450 capable of catalysing both hydroxylation and epoxidation reactions at the isoprenyl tail (C17–C20) of mangicdiene (Fig. 5c,f). Mangicol E was detected in strain AO-Y55 (containing MgcCDE), suggesting that mgcC can catalyse the ring-opening hydrolysis of the C19, C20-epoxide of mangicol J (Fig. 5d), and thus is an epoxide hydrolase; this result was further confirmed in vitro (Fig. 5e and Extended Data Fig. 7). Mangicol H was detected in strain AO-Y52 (containing MgcDE), suggesting that a native epoxide hydrolase of AO (an isozyme of MgcC) can catalyse the ring opening of **A** to **B** and finally to **68** (Fig. 5c,f). Mangicols K and L, derived from mangicol E via C7 and C8 hydroxylation, respectively, were detected in strain AO-Y57 (containing MgcCDEF), suggesting the P450 mgcF acts on the core structure (Fig. 5d).

These combined results indicate a putative biosynthetic pathway for mangicols H–L, shown in Fig. 5f. In this proposal, MgcD catalyses the condensation of isopentenyl pyrophosphate and dimethylallyl

pyrophosphate into geranylgeranyl pyrophosphate, and then the CI–III–IV cyclization to **64**. MgcE catalyses the C18, C19-epoxy substitution of **64** to make **A**, and MgcC hydrolyses this epoxide to form **B**; MgcE then catalyses the oxidation of the C18-hydroxyl to yield **68**. Alternatively, MgcE can sequentially catalyse C18-hydroxylation to form **C** followed by C19, C20-epoxy substitution to make **D**, C18-carbonylation to form **69**, and C17-hydroxylation to make **70**. MgcC can then hydrolyse the C19, C20-epoxide to yield **71**. Finally, MgcF catalyses either C7 or C8 hydroxylation to give **72** or **73**, respectively. This proposal presumes that a native epoxide hydrolase of AO can catalyse ring-opening en route to **68** in strains AO-Y52 and AO-Y56. The auto-HTP biofoundry workflow thus makes it possible to quickly identify both metabolic intermediates and final products, and rationalize the corresponding biosynthetic mechanism, without the multiple rounds of gene activation, knock-outs and knock-ins that are required in traditional homologous and heterologous expression strategy.

AO chassis 2.0 for overproduction of mangicol J. The AO chassis 1.0 served as an ideal starting point to mine diverse BGCs and undertake initial explorations of bioactivity. However, detailed investigations of specific compounds require more efficient production; for example, strains AO-Y51 and AO-Y52 only accumulate 0.66 mg l⁻¹ of mangicdiene and 0.08 mg l⁻¹ of mangicol J, respectively, which is insufficient for follow-up studies. Previous research has indicated that optimizing the mevalonate (MVA) pathway can obviously enhance the titre of terpenoids, and that tHMG1 and Idi are the rate limiting enzymes⁴⁹. Thus, to amplify compound production, we systematically engineered the MVA pathway of AO to establish an efficient chassis 2.0 for the overproduction of mangicdiene and mangicol J (Fig. 6 and Extended Data Figs. 8 and 9). We started by creating strains with improved production of the mangicdiene core scaffold. We constructed plasmids harbouring the entire endogenous MVA pathway, an additional three copies of *tHMG1* and *mgcD*, and randomly inserted them into the chromosome of AO NSAR1 to produce strains AO-S81 to AO-S84. In parallel, on the basis of a previously developed CRISPR–Cas9 method³⁰, we developed a CRISPR–Cas9-mediated site-specific integration

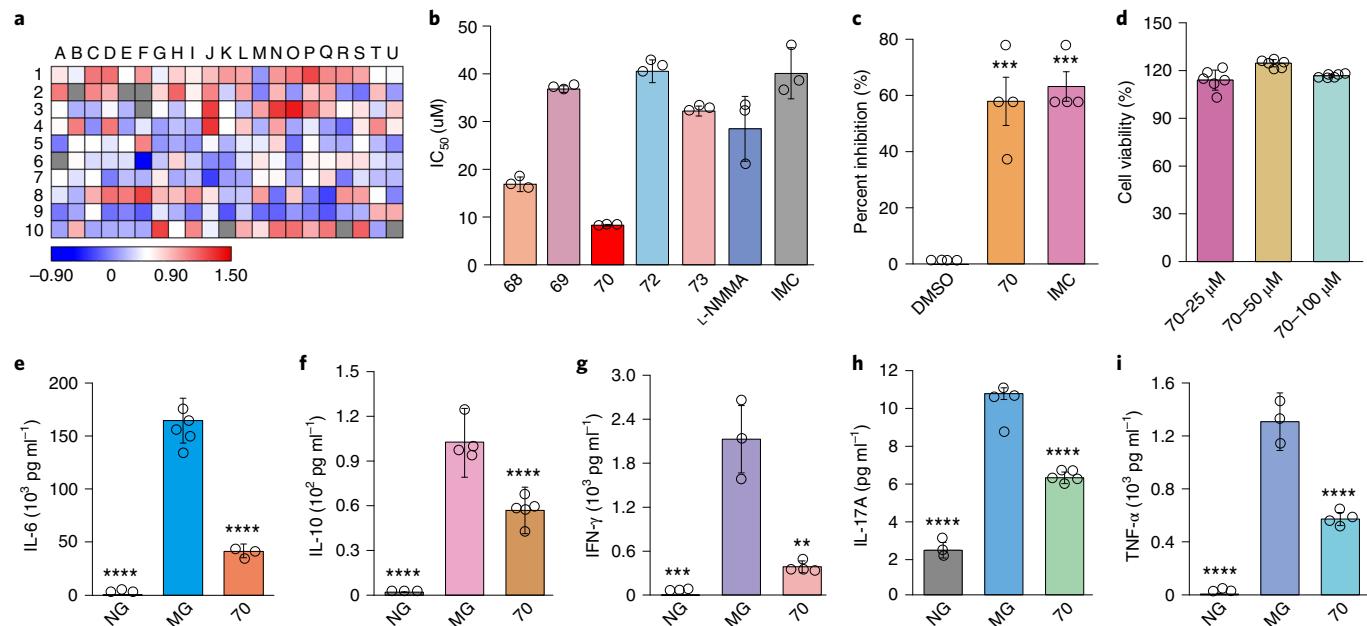


Fig. 4 | Anti-inflammatory activity screening of terpenoids from BGC11-FgMS. **a**, Anti-inflammatory activity of extracted crudes produced by AO-strains. The scale shows the effects of compounds on nitric oxide (NO) production in the LPS-stimulated murine macrophage cell line RAW264.7. The macrophage cells were treated with extracted crudes produced by AO-strains at a final concentration of 500 $\mu\text{g ml}^{-1}$. NO was detected with Griess reagents at 540 nm using a microplate reader. T10 is a negative control, containing products extracted from the wild-type AO-strain. The data are presented as the average of three parallel samples. **b**, Compounds produced by the FgMS gene cluster inhibited LPS-induced NO secretion in RAW 264.7 macrophages. Cells were treated with mangicols H-L (compounds **68–70**, **72** and **73**) and positive controls (l-NMMA , IMC) at different concentrations to calculate the IC_{50} value. Data are expressed as mean \pm s.d. of three independent experiments. l-NMMA : N^{G} -monomethyl- L -arginine, monoacetate salt (inhibitor of nitric oxide synthase); IMC is a non-steroidal anti-inflammatory agent. **c**, Evaluation of anti-inflammatory activity of mangicol **70** in ear oedema mice. The model mice were treated with **70** and IMC at the same dosage of 100 μg per ear. The IMC group represented the positive control. The data are expressed as mean \pm s.e.m. ($n=4$). Statistically significant differences relative to the control group (DMSO) was performed using a two-tailed Student's t-test ($^{**}P < 0.01$ and $^{***}P < 0.005$). $P^{\text{70}} = 0.00121$, $P^{\text{IMC}} = 0.000147$. **d**, Effects of cell viability of mangicol **70** on murine macrophage cell line RAW264.7. The data are presented as mean \pm s.d. ($n=6$). **e–i**, Effects of mangicol **70** on production of the inflammatory cytokines interleukin-6 (IL-6) (**e**), interleukin-10 (IL-10) (**f**), interferon- γ (IFN- γ) (**g**), interleukin-17A (IL-17A) (**h**) and tumour necrosis factor (TNF- α) (**i**) in LPS-stimulated mice. The inflammatory mice were treated with mangicol **70** at the dosage of 10 mg kg^{-1} (10 mpk), respectively. **e–i**, The data are presented as mean \pm s.d. (NG, $n=3$; MG and **70**, $n=5$). NG, negative group, treated only with saline; MG, model group, treated only with LPS. Statistical analysis was carried out using a two-tailed Student's t-test ($^{**}P < 0.01$, $^{***}P < 0.005$, $^{****}P < 0.001$). P values in **e**, $P^{\text{NG}} = 3.69 \times 10^{-6}$, $P^{\text{70}} = 2.58 \times 10^{-6}$. P values in **f**, $P^{\text{NG}} = 2.72 \times 10^{-5}$, $P^{\text{70}} = 2.41 \times 10^{-4}$. P values in **g**, $P^{\text{NG}} = 2.74 \times 10^{-3}$, $P^{\text{70}} = 6.00 \times 10^{-3}$. P values in **h**, $P^{\text{NG}} = 1.82 \times 10^{-5}$, $P^{\text{70}} = 1.70 \times 10^{-6}$. P values in **i**, $P^{\text{NG}} = 1.62 \times 10^{-4}$, $P^{\text{70}} = 3.52 \times 10^{-4}$.

system and integrated the 14-kb cassette (fragment from pSC247) via homologous recombination with 85% efficiency (Fig. 6a and Supplementary Table 9). For site-specific integration, we used the high expression loci (hot spots (HS), Extended Data Fig. 10) HS201, HS401, HS601 and HS801 as our target sites to promote the expression of functional genes³⁰. Specifically, we integrated the entire MVA pathway, and additional four copies of *tHMG1* at a different HS site of AO to generate AO-S95. We then integrated one copy of *mgcD* into the HS801 site of AO-S95 to produce AO-S96. Compared with the parent strain AO-Y51, the titre of mangidiene were increased 41-fold from 0.66 mg l^{-1} to 27.38 mg l^{-1} in AO-S96 and 133-fold to 87.84 mg l^{-1} in AO-S84 (Fig. 6b,c and Extended Data Fig. 9). These titres are obviously higher than those of the sesquiterpenes detected in *E. coli* and *S. cerevisiae*^{24,50}, providing an ideal platform for the high-throughput genome mining of terpenoid BGCs of filamentous fungal origin.

Next, we created strains for the overproduction of mangicol J using random insertion and site-specific integration of functional genes. For random insertion, two copies of *mgeE* were randomly inserted into the chromosome of AO-S84 to produce strain AO-S94. For site-specific integration, we integrated one copy of *mgcD* and two copies of *mgeE* into the HS801 site of AO-S95 to produce AO-S97. Compared with strain AO-Y52, the titre of mangicol J was increased

by 112-fold from 0.08 to 8.93 mg l^{-1} in AO-S98 and by 151-fold to 12.09 mg l^{-1} in AO-S94 (Fig. 6c,d and Extended Data Fig. 9).

Finally, we deleted *mgeD* from strain AO-S84 to produce AO-S85. This strain thus serves as a universal chassis 2.0 to characterize genes and terpenoid structures, providing abundant precursors for the overproduction of fungal specific terpenoids; the HS-specific integrating strain AO-S95 serves as powerful complement for stable overproduction of fungal derived terpenoids. We confirmed the generality of this advanced chassis by refactoring the BGC37-Tvi09626 gene cluster into the AO-S95 strain. Compared with chassis 1.0, the titre of oxygenated and glycosylated products was obviously improved (Supplementary Fig. 37d).

Discussion

Genome mining to identify new bioactive compounds is an important endeavour, both to improve our understanding of the natural world and to provide new resources for development of pharmaceutical agents. The classic homologous and heterologous expression strategy has provided a wealth of molecules and information, but can be limited by the number of manual steps involved in the genetic manipulations, as well as challenges in expressing membrane-localized CYP450s and proper recognition of intron-rich genes. While efforts have been made to develop more

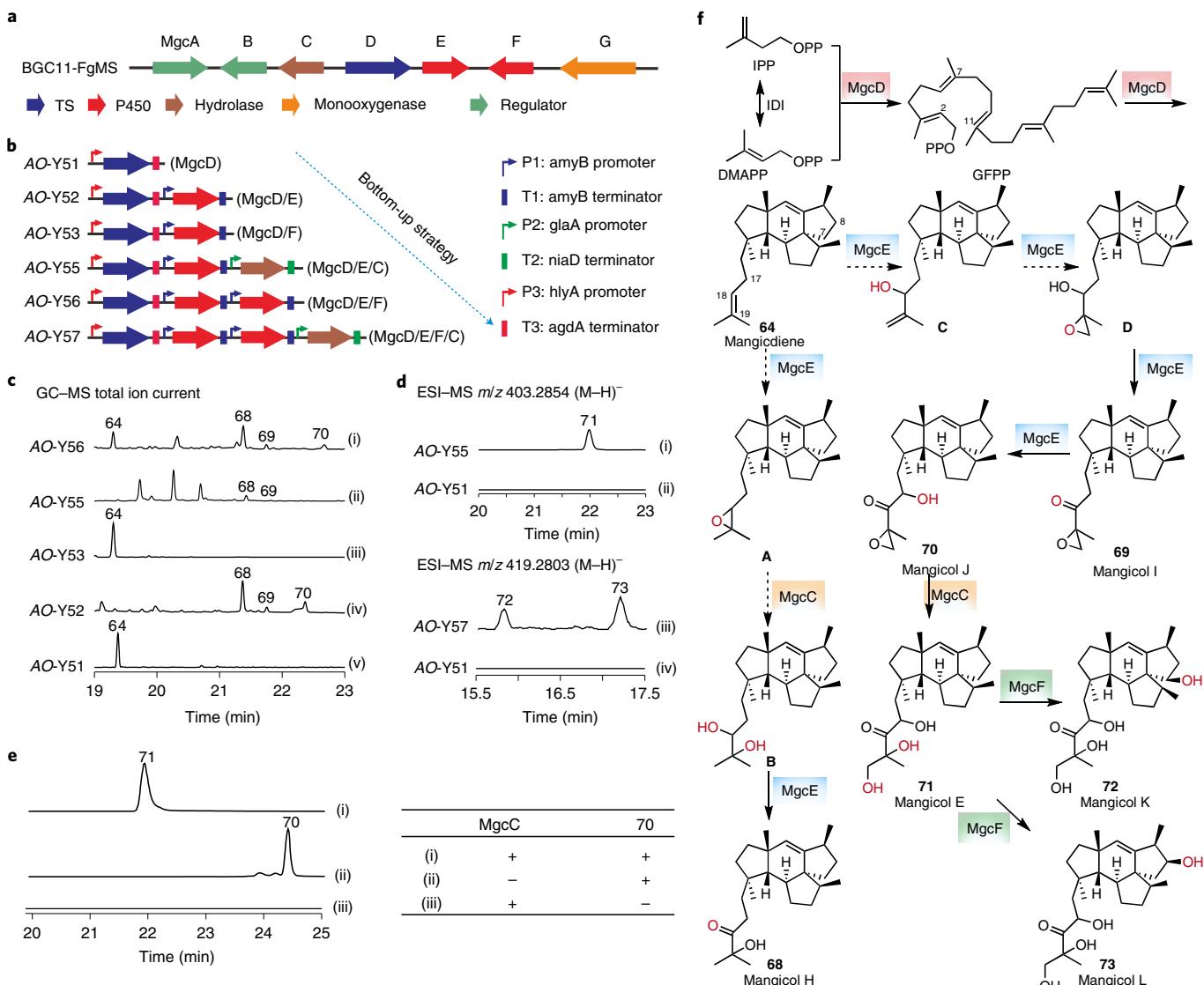


Fig. 5 | Mechanistic characterization of discovered mangicols from BGC11-FgMS. **a**, Bioinformatic analysis of the FgMS gene cluster. **b**, Reconstitution of the FgMS gene cluster in *A. oryzae*. **c**, GC-MS analysis of compounds produced by AO-strains expressing the FgMS gene cluster. AO-Y51 (v) contains the MgcD enzyme (terpene synthase, TS); AO-Y52 (iv) contains MgcD and the cytochrome P450 MgcE; AO-Y53 (iii) contains MgcD and the cytochrome P450 MgcF; AO-Y55 (ii) contains MgcD, MgcE and the epoxide hydrolase MgcC; AO-Y56 (i) contains MgcD, MgcE and MgcF. **d**, HR-ESI-MS analysis of products produced by AO-strains expressing the FgMS gene cluster. AO-Y57 (iii) contains MgcD, MgcE, MgcC and MgcF. **e**, Functional characterization of MgcC in vitro. Compound **71** was detected using HR-ESI-MS in the reaction containing MgcC and substrate **70** (i). The reactions without MgcC (ii) or without substrate (iii) did not produce **71**. **f**, Proposed biosynthetic pathway for the mangicols produced by the FgMS gene cluster.

universal expression platforms, there are none intended for filamentous fungi; we anticipated that a chassis built for this purpose would facilitate the discovery of products from closely related species. Herein, we developed the auto-HTP biofoundry platform to enable efficient and parallel genome mining of terpenoids in AO, as well as subsequent screening and biosynthetic pathway analysis.

Using the auto-HTP biofoundry platform, we reconstituted 39 BGCs into 208 AO-strains and characterized 185 terpenoids in a short period of time, and the overall strategy means that the size of the refactored BGCs and the resulting number of strains can easily be scaled up (Table 1). The refactored clusters enabled us to identify functionally important intermediates, such as mangicol J, that could easily have been missed using a traditional heterologous expression approach. The ability of mangicol J to decrease the levels of inflammatory cytokines makes it an attractive candidate for the development of anti-inflammatory agents as well as for investigating

the anti-inflammatory mechanism in mice and humans. The construction of the AO chassis 2.0 helped us increase the titre of mangicol J by 151-fold to 12.09 mg l^{-1} , enabling further research on this intriguing molecule. Moreover, the chassis 2.0 provides a universal platform for the efficient genome mining and overproduction of terpenoids from filamentous fungi, enabling researchers to find other new molecules that may affect drug discovery.

Nevertheless, there are limitations to our strategy. First, the auto-HTP biofoundry pipeline requires several pieces of equipment that must communicate smoothly. Second, the BGCs of interest must be informatically well defined. Third, the efficiency of the auto-HTP platform does not extend to compound purification and structural characterization. In our case, the structures of a large number of terpenoids discovered in this work are currently unidentified. It is also worth noting that the refactored strains corresponding to 13 of the 39 clusters did not produce detectable

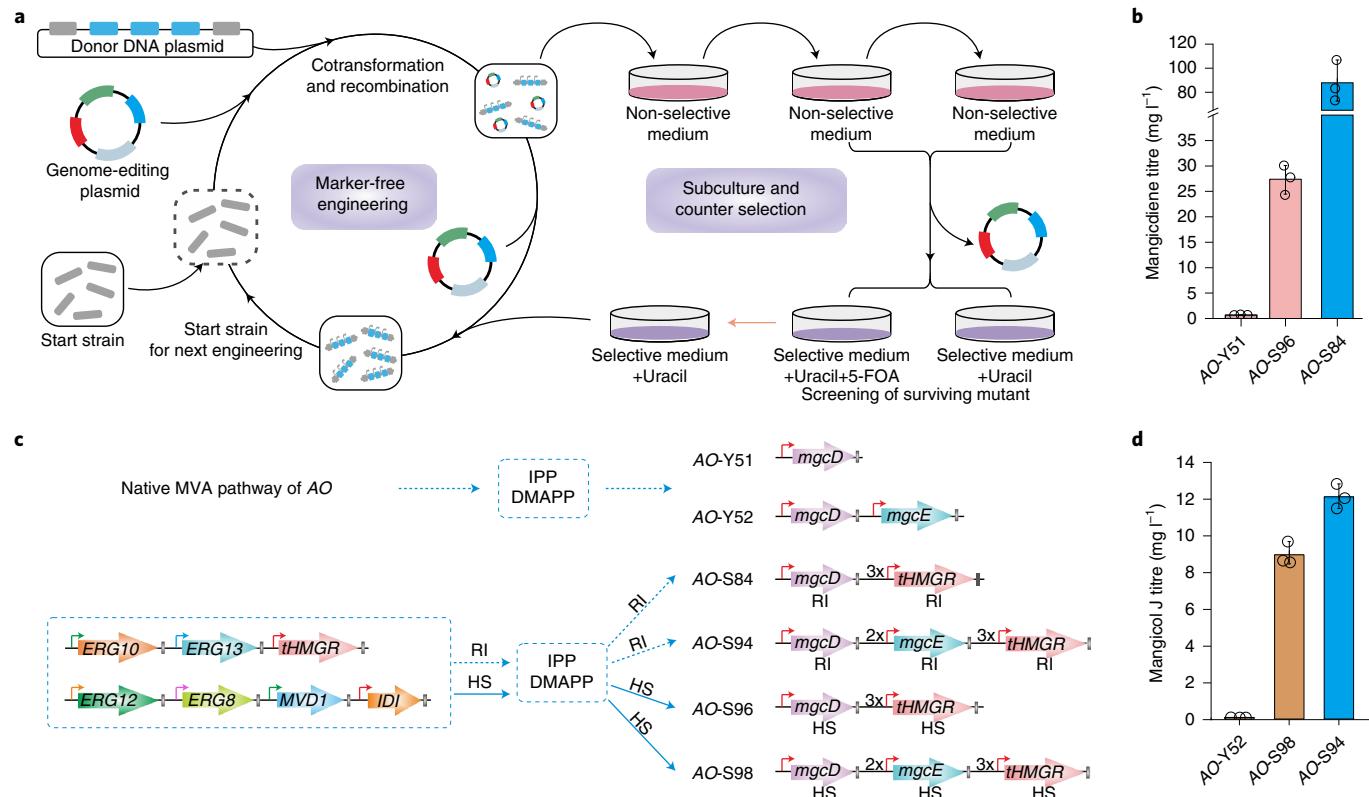


Fig. 6 | AO chassis 2.0 for overproduction of mangicdiene and mangicol J. **a**, The entire workflow of the CRISPR-Cas9-mediated marker-free recombineering system. **b**, Production of mangicdiene in engineered AO-strains. Data are presented as mean \pm s.d. ($n=3$). **c**, The strain for production of mangicdiene and mangicol J via the endogenous MVA or overexpressed MVA pathways. **d**, Production of mangicol J in engineered AO-strains. Data are presented as mean \pm s.d. of ($n=3$). RI, random insertion; AO-Y51: *mgcD*, random insertion; AO-S96: *mgcD*, MVA pathway gene, 3x*HMG1*, HS integration; AO-S84: *mgcD*, MVA pathway gene, 3x*HMG1*, random insertion; AO-Y52: *mgcD*, *mgeE*, random insertion; AO-S98: *mgcD*, 2x*mgeE*, MVA pathway gene, 3x*HMG1*, HS integration; AO-S94: *mgcD*, 2x*mgeE*, MVA pathway gene, 3x*HMG1*, random insertion.

terpenoids. We suspect this is mainly due to evolutionarily dead terpene synthases; more insights into this enzymatic family might facilitate more robust BGC selections in the future. Of course, there are also continued challenges when undertaking expression in a heterologous host; substrates for some of the enzymes (for example, the rare glycosyltransferases) may not be present, leading to a false negative and efforts to reconstruct biosynthetic pathways must take into account that the native enzymes can interfere with the intended reaction sequence. Further development of our AO chassis may be able to address these issues.

In summary, our auto-HTP biofoundry platform-based genome mining strategy in AO enables efficient genome mining and determination of biosynthetic mechanisms for bioactive terpenoids from filamentous fungi. Our advanced AO chassis 2.0 provides a stable platform for the improved production of fungus-derived terpenoids and increases the chances of mining terpenoids that are produced via unusual catalytic mechanisms. Both these approaches can be easily adapted to other natural products and microorganisms, which might accelerate the discovery and development of new pharmaceutical agents.

Methods

Materials and reagents. PCR amplification for plasmid construction was performed using Phusion High-Fidelity polymerase (New England Biolabs, NEB) and Prime STAR GXL DNA polymerase (Takara Bio, Inc.), except for colony PCR, for which 2x Taq Plus Master Mix (Vazyme) was used. The PCR primers were synthesized using GenScript. Fast digestion restriction enzymes were purchased from Thermo Fisher Scientific. Enzymes for Goldengate and USER cloning were purchased from NEB. Cloning was performed using the chemically competent *E. coli* strain DH10B (Invitrogen). All yeast strains were cultured in selective

uracil-dropout media. AO transformants were selected under triple auxotrophic culture (arg-, ade- and met-) conditions. Yeast extracts and tryptone used for preparing Luria-Bertani (LB) medium were purchased from Oxoid. All salts and reagents were purchased from Sinopharm Chemical Reagent Co., Ltd. The animal care and use were adhered to the Chinese National Guidelines for Ethical Review of Animal Welfare. Animals were handled according to the Guidelines of the China Animal Welfare Legislation, and the study protocol was approved by the Institutional Animal Care and Use Committee (IACUC) of Renmin Hospital of Wuhan University Human Research Ethics Committee (IACUC issue no. 20201220A).

Preparation of fragments for automated plasmid construction. Functional genes were amplified from the genomes of the five filamentous fungi. For each gene, 40 base pairs overlapping sequences were added to the 5' and 3' terminals, corresponding to the specific promoter and terminator, respectively. Regulator cassettes containing promoters and terminators were amplified from the plasmids described in Supplementary Table 12. The primers used in this study are listed in Supplementary Table 11. The optimized vector backbones (pTAex3-, pUSA- and pAdeA-optimized) were constructed by inserting the amplification and screening cassettes of *S. cerevisiae* into the original vectors of pTAex3, pUSA and pAdeA³⁴, respectively. Except for the amplification and screening cassettes of *E. coli* (pBR322_ori and ampicillin resistance gene) and *S. cerevisiae* (CEN_ori and uracil synthesis gene), the pTAex3-, pUSA- and pAdeA-optimized vector backbones contained the screening cassettes Arg, Sc and Ade of AO, respectively (Extended Data Fig. 2). The vector backbones containing three types of amplification and screening cassettes enabled the plasmids to be shuttled between *E. coli*, *S. cerevisiae* and AO. All amplified fragments were purified using VAHTS DNA Clean Beads (N411-01, Vazyme).

Plasmid construction using an automated yeast assembly method. For high-throughput assembly of plasmids, the LiAc/ss carrier DNA/polyethylene glycol (PEG) yeast transformation protocols were modified as previously reported³⁵. *S. cerevisiae* CEN.PK2-1D (EUROSCARF) with the genotype *MAT**a**ura3-52trp1-289leu2-3 112his3Δ1 MAL2-8C SUC2* was used for all yeast heterologous recombination experiments. For each assembled plasmid, 300 ng

of each coding sequence fragment was combined with 300 ng of each required regulator cassette and 300 ng of the appropriate linearized expression vector. Using the Biomek FX^p Laboratory Automation Workstation (Beckman Coulter) equipped with an MP200 96-Tip Tool for liquid-handling operations, this DNA mix was transformed into the CEN.PK2-1D strain by using the LiAc/PEG protocol. Detailed procedures are presented in Supplementary Fig. 2. Biomek Software runs the robotics platform based on the specified programs for various steps along the yeast assembly cycle, such as adding DNA mixture combined with LiAc/PEG solution into competent cells for transformation. Each program outlines a series of basic functions that do not change between assembly sets, such as 'load P200 tips', 'aspirate', 'mix' or 'shake'. Detailed information about constructing plasmids is provided in Supplementary Table 12.

E. coli transformation for plasmid enrichment and verification. *E. coli* DH10B competent cells were dispensed into 96-well deep plates at 70 µl per well, mixed with 10 µl of yeast plasmids and incubated at 4 °C for 30 min. The plates were then transferred into an incubator preheated to 42 °C and incubated for 3 min; next, 0.8 ml of the LB medium was added to each well and incubated at 37 °C for 45 min. Cell cultures were harvested by centrifugation at 1,500g for 8 min and the supernatants were discarded. The remaining 50 µl of the cell culture was transferred onto LA medium containing the appropriate antibiotic. The colonies were picked with Molecular Device Qpix 460 and inoculated into LB medium for plasmid enrichment. The plasmids obtained were further verified using the appropriate restriction enzymes. Detailed procedures are shown in Supplementary Fig. 3.

Plasmid extraction. Yeast colonies were placed in 1.5 ml of uracil-dropout medium, seeded onto 96-well deep-well plates and incubated with shaking at 998 r.p.m. and 30 °C for 16–18 h. Cell cultures were harvested by centrifugation at 3,500 r.p.m. for 8 min, and the supernatants were discarded. Plasmid DNA was extracted using the magnetic bead method with a modified version of the MagPure Plasmid LQ Kit (Magen Bio). Lyticase (500 U ml⁻¹, Sigma-Aldrich, 20210108) combined with 250 µl of buffer S1 was added to the cell culture and incubated at 25 °C for 2 h. For the detailed procedures, see Supplementary Fig. 4.

Construction of AO-strains for fungal BGC expression. We selected AO NSAR1 (*niaD*⁻, *sC*⁻, *ΔargB*, *adeA*⁻) as the heterologous expression host, and used the protoplast-polyethylene glycol method for protoplast transformation. The mycelium with spores was inoculated in 100 ml of dextrin-peptone-yeast liquid medium (2% dextrin, 1% polypeptone, 0.5% yeast extract, 0.5% KH₂PO₄, 0.05% MgSO₄·7H₂O), and cultivated for 2 d at 30 °C and shaking at 140 r.p.m. The protoplasts were obtained by digestion with yatalase (TaKaRa, 2.0 mg ml⁻¹) and lysing enzymes (sigma, 3.0 mg ml⁻¹).

For the high-throughput transformation of AO, the following liquid-handling operations were performed using the Biomek FX^p Laboratory Automation Workstation (Beckman Coulter) equipped with MP200 96-Tip Tool and flexible 8 channels. Plasmid mixtures (10–15 µg each) were automatically aspirated using the P200 tips at the speed of 100 and dispensed into 96-well deep-well plates at 2 mm from the bottom of the wells, in which containing 100 µl of protoplast suspension that was picked in advance by the Thermo Scientific Multidrop Combi SMART Dispenser. After mixing five times at 100 µl s⁻¹, the aliquots were incubated at 4 °C for 1 h, following which 1.25 ml of PEG solution (25% PEG 6000, 100 mM CaCl₂, 0.6 M KCl, 10 mM Tris-HCl, pH 7.5) was automatically aspirated using P200 tips at a low speed of 50 and dispensed into the pellets at a speed of 100, as well as set stick wall mode of tips to reduce the loss of the ropy PEG solution. After incubation at 25 °C for 30 min, the mixture was centrifuged at 420g for 25 min. The supernatant was partially discarded, and the resulting 200 µl of solution was diluted with 1 ml of STC buffer (1.2 M Sorbitol, 10 mM CaCl₂, 10 mM Tris-HCl, pH 7.5). After centrifugation at 420g for 25 min, the supernatant was partially discarded resulting about 100 µl of the mixture that was automatically aspirated at the speed of 100 using flexible 8 channels with P200 tips, dispensed at 10 mm above the wells and coated with shaking at 700 r.p.m. on a triple auxotrophic (*arg*⁻, *ade*⁻ and *met*⁻) solid medium (2% glucose, 1.2 M sorbitol, 0.2% NH₄Cl, 0.1% (NH₄)₂SO₄, 0.05% KCl, 0.05% MgSO₄·7H₂O, 0.15% KH₂PO₄, 1.6% agar, pH 5.5) in 24-well deep plates (5 ml of medium per well); the plates were incubated at 30 °C for 2–5 d. Detailed procedures are shown in Supplementary Fig. 5.

For the verification of AO transformants, 0.1–1 µg mycelia of AO colonies were picked manually and suspended in 100 µl of NaOH solution (25 mM) in 96-well PCR plates and lysed by heating at 100 °C for 10 min in a PCR machine for genomic DNA extraction. The supernatant of the mixture was used as a template for PCR verification using 2×Taq Plus Master Mix (Vazyme). All the constructed AO-strains are listed in Supplementary Table 13.

Fermentation and extraction of terpenoids. Authenticated AO transformants were initially cultured on dextrin-peptone-yeast agar plates at 30 °C for 5 d to obtain fresh mycelia and spores, which were then inoculated in 24-well deep plates containing solid medium (1 g of rice and 1.5 ml of distilled H₂O per well) for product detection. Alternatively, they were inoculated in plastic lunch boxes containing rice medium (120 g of rice and 100 ml of distilled H₂O per box) for product structure elucidation. Each culture was incubated at 30 °C for 2 weeks.

Fermented cultures in 24-well deep-well plates were extracted twice by successively soaking in acetone and ethyl acetate (5 ml per well) at room temperature for 2 h each. The organic solvent was combined and evaporated in a fume hood to obtain crude extracts (10–20 mg per well). Fermented cultures in a plastic lunch box were extracted three times using ethyl acetate. The organic phases were combined and concentrated to dryness using a rotary evaporator.

In vitro and in vivo anti-inflammatory assays of terpenoids. Inhibition of NO production by LPS-stimulated RAW 264.7 murine macrophages purchased from the American Type Culture Collection (ATCC, TIB-71) was used to evaluate the functions of the extracted terpenoids. The cells (2.0 × 10⁵ cells per well) were seeded onto 24-well plates and treated with 1 µg ml⁻¹ LPS in the absence or presence of the tested products for 16 h. Then, 50 µl of supernatant was transferred to a new 96-well plate, and 50 µl of Griess reagent I and 50 µl of Griess reagent II (Nitric Oxide Assay Kit, S0021, Beyotime) were added sequentially. Absorbance was measured at 540 nm on an EnSpire Multimode Plate Reader. The procedure details are shown in Supplementary Fig. 6. IMC and L-NMMA were used as positive controls.

The cytotoxicity of mangicol compounds was tested on RAW 264.7 cells using the Cell Counting Kit-8 (CCK-8) (Dojindo) method as follows. The exponentially growing cells were seeded onto 96-well plates at a density of 2.0 × 10⁴ cells per well and incubated for 24 h at 37 °C in 5% CO₂. The tested products at the indicated concentrations were added to each well and incubated for 24 h. Subsequently, the CCK-8 solution (10 µl) was added to each cell and incubated for another 1 h. The absorbance was measured at 450 nm using an EnSpire Multimode Plate Reader. Wells without drugs were used as blank controls.

Eight-week-old female ICR (CD-1) mice (Harlan Sprague Dawley Inc.) were divided into three groups with four mice per group, housed with 12 h:12 h dark/light cycle, temperature (22–25 °C), humidity (50–60%), provided with food and water ad libitum and acclimatized for 1 week. Each mouse received 2 µg per ear phorbol-12-myristate-13-acetate (PMA, 2 µg in 20 µl of acetone) in the right ear. Initially, the right ears were treated with DMSO (20 µl, vehicle control), IMC (100 µg in 20 µl of DMSO) or mangicol J (100 µg in 20 µl of DMSO), while the left ears received vehicle (20 µl of DMSO). One hour later, the phlogistic agent was applied using an automatic pipette in 10 µl volumes to both the inner and outer surfaces of the right ear. The left ear (control) received the vehicle (20 µl of acetone). Ear oedema was measured at 6 h post-PMA treatment using a digital caliper and calculated by subtracting the thickness of the left ear from the right ear. The data obtained are expressed as the mean ± s.d.; the unpaired Student's *t*-test was used to determine statistical significance.

The anti-inflammatory activity of mangicol J based on an acute inflammation model was tested in mice using flow cytometry. Seven-week-old male C57BL/6J mice (Zhejiang Weitong Lihua Experimental Animal Technology Co., Ltd) were divided into three groups (three mice in the negative group, and five mice each in the model and experimental groups), housed with 12 h:12 h dark/light cycle, temperature (22–25 °C), humidity (50–60%), provided with food and water ad libitum and acclimatized for 1 week. The mice were stimulated with an intraperitoneal injection of 20 mg kg⁻¹ LPS. After 1 h, the mice in which inflammation was induced were treated with compound 70 by intraperitoneal injection at a dose of 10 mg kg⁻¹. Blood samples were collected after 6 h and analysed; the presence of interleukin (IL)-6, IL-10, IFN-γ, IL-17A and TNF-α was detected using a BD FACSCalibur^M Flow Cytometer by FCAP Array Software using a Mouse Th1/Th2/Th17 CBA Kit (560485; BD Biosciences).

Inhibition of p-STAT3 protein level in JAK2-STAT3 signalling by IL-6 and interleukin-6 receptor subunit alpha (IL-6Rα)-stimulated human umbilical vein endothelial cells purchased from the ATCC (PCS-100-010) was used to evaluate the anti-inflammatory activity of mangicol J. The cells (2.0 × 10⁵ cells per well) were seeded onto 12-well plates at a density of 2.0 × 10⁴ cells per well and incubated for 24 h at 37 °C in 5% CO₂. The negative group did not have treatment with IL-6 and IL-6Rα. The model group was treated with IL-6 (5 ng ml⁻¹) and IL-6Rα (25 ng ml⁻¹) simultaneously. The experimental group was treated with compound 70 (2 µg ml⁻¹) for 3, 6 and 12 h after treatment with IL-6 (5 ng ml⁻¹) and IL-6Rα (25 ng ml⁻¹). Cells from each group were respectively gathered and lysed for cleavage protein for western blotting to analyse the concentration of β-actin, STAT3 and p-STAT3 proteins. The antibodies used are as follows, first antibody: β-Actin (mouse source, catalogue no. ab8226, dilution ratio, 1:10,000, abcam), p-STAT3 (rabbit source, catalogue no. 9145s, dilution ratio, 1:1,000, CST) and STAT3 (rabbit source, catalogue no. 12640s, dilution ratio, 1:1,000, CST); second antibody: hRP-goat anti-rabbit (goat source, catalogue no. E-AB-1003, dilution ratio, 1:3,000, Elabscience) and hRP-goat anti-mouse (goat source, catalogue no. E-AB-1001, dilution ratio, 1:3,000, Elabscience). The relative expression levels of the target proteins were calculated by dividing the intensity ratio of the target proteins by that of STAT3. The data obtained are presented as mean values ± s.d. and the unpaired Student's *t*-test was used to determine statistical significance.

Plasmids for overproduction of mangicdiene and mangicol J. To overproduce mangicdiene, we constructed a series of plasmids to overexpress the MVA pathway genes. MVA pathway genes, promoters and terminators (Supplementary Table 10) were amplified using the corresponding primers from AO complementary DNA

listed in Supplementary Table 14. A series of plasmids (Supplementary Table 15) was constructed using the yeast assembly method. To obtain additional copies of tHMG1 and FgMS, PhlyA-tHMG1-Tnos was amplified from pSC44 using the primer pair 59-1F/59-1R and inserted into SacI-digested pSC44 to generate pSC59. To overexpress mangicol J, pSC111 was constructed using a yeast assembly. To obtain an additional copy of *mgcE*, PhlyA-mgcE-Tnos was amplified from pSC111 using the primer pair 112-1 NotI F/112-2 HpaI PacI R and inserted into NotI/PacI-digested pSC111 to generate pSC112.

Plasmids for Cas9 genome editing. To achieve a high efficiency of CRISPR-Cas9-mediated homologous integration, we systematically tested and optimized the CRISPR-Cas9 method and constructed a series of plasmids. Specific genomic RNA sequences targeting *pyrG*, *HS201*, *HS401*, *HS601* and *HS801* were obtained using the open-source tool <http://www.biotools.com/index.html> (Extended Data Fig. 10)⁵¹. A series of plasmids were constructed to evaluate the influence of the number of single-guide RNAs, the introduction of AMA1 sequence⁵² and the concentration of sgRNAs towards genome-editing efficiency. The yeast replication element, *CEN*, was selected as an arbitrary sequence, amplified from pSC01 using the primer pair 78-1F/78-1R and inserted into XbaI/BamHI-digested pET28a to generate pSC78. Next, using the Goldengate method, sgRNA expression cassettes were digested with *Bsa*I and inserted into pSC78 to generate the plasmids pSC93-pSC96, pSC98 and pSC252. *P_{AfU3}*-transfer RNA-ligD sgRNA-tRNA-*T_{AfU3}* was synthesized and inserted into pET28 to generate pSC92. The Pvul/NotI-digested fragments of pSC92, pSC93 and pSC94 were inserted into pSC121(PacI/NotI-digested) to generate pSC131, pSC132 and pSC133, respectively. The PacI/NotI-digested fragments of pSC92, pSC93 and pSC94 were inserted into pSC60 (PacI/NotI-digested) to generate pSC65, pSC69 and pSC70, respectively. To assess the *P_{U6}*-based CRISPR-Cas9 system, 96F1 was amplified from the NSAR1 genome by primer pair 96-1F/96-R, 116F2 was amplified from pUCm-gRNA scaffold-eGFP by primer pair 96-2F/116-2R and 116F3 was amplified from the NSAR1 genome using primer pair 116-3F/116-3R; three fragments were assembled using primer pair 96-1F/116-3R through overlap extension-PCR (OE-PCR) and inserted into pSC109 (NotI/PacI-digested) to generate pSC116; 117F2 was amplified from pSC116 by primer pair 96-2F/117-2R; 96F1 and 117F2 were assembled by OE-PCR; 117F3 and 117F4 were amplified from pSC116 by primer pairs 117-3F/117-3R and 117-4F/116-3R, respectively, and then assembled by OE-PCR. The two resulting PCR fragments NotI/BcuI and BcuI/PacI were digested and inserted into pSC109 (NotI/PacI-digested) to generate pSC117. The pSC98 NotI/PacI-digested fragment 3x P_{U6} -*pyrG* sgRNA-*T_{U6}* was inserted into pSC109 (NotI/PacI-digested) to generate pSC118.

The pSC98 NotI/SmaI-digested fragment 3x P_{U6} -*pyrG* sgRNA-*T_{U6}* was inserted into pSC87 (NotI/SmaI digested) to generate pSC184. To knock down the *FgMS* gene in strain AO-S84, pSC252 NotI/PacI-digested fragment 3x P_{U6} -*FgMS* sgRNA-*T_{U6}* was inserted into pSC251 (NotI/PacI-digested) to generate pSC253. To overproduce mangicol J, 24-F1 and F2 were assembled using the primer pair 24-1F/254-2R. Then, PacI/NotI-digested 24-F1 + F2 was inserted into pSC134 (NotI/PacI-digested) to generate pSC24.

Strains for the overproduction of mangicol J in AO. To build a more efficient AO chassis for genome mining and overproduction of mangicol J, the native MVA pathway and the downstream mangicol J-forming pathway were engineered under the control of the desired promoters. Finally, a series of AO-strains were constructed, which are listed in Supplementary Table 16. AO-S84 was constructed by transforming the corresponding plasmid into AO NSAR1. AO-S184 was constructed by transforming pSC184 into AO NSAR1. AO-S95 was constructed using three genome-editing cycles. First, pSC249 and pSC246 were cotransformed into AO-S184, and the constructed strains were PCR-confirmed to contain inserted flanks at the HS401 site. After two or three subcultivations under non-selective conditions, the conidial suspension (10⁴ conidia in 5 µl) of the strain was spotted onto a plate containing uracil and 5-FOA and a plate containing uracil without 5-FOA; strains that can grow under the two culture conditions were screened to serve as the starting strain for the next round of gene editing. The plasmids used in the second and third rounds of gene editing were pSC263, pSC260, pSC248 and pSC262. The construction of AO-S96 to AO-S98 is described in Supplementary Table 16. The engineered plasmids and strains were verified by PCR amplification to ensure that all the plasmids were correctly constructed and all the genes were correctly integrated into the genome of AO, as required.

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

Data availability

Data supporting the findings of this study are available within the article and its Supplementary Information files. The accession numbers (NMDCN0000QH2 to NMDCN0000QH9, NMDCN0000QHA to NMDCN0000QHV, and NMDCN0000QI0 to NMDCN0000QI9) and nucleotide sequences for the characterized enzymes were deposited in the National Microbiology Data Center (<https://nmdc.cn/en>) will be available after publication. Source data are provided

with this paper. All other data that support the findings of this study are available from the corresponding author upon reasonable request.

Received: 4 June 2021; Accepted: 4 March 2022;

Published online: 26 April 2022

References

- Wright, G. D. Opportunities for natural products in 21st century antibiotic discovery. *Nat. Prod. Rep.* **34**, 694–701 (2017).
- Newman, D. J. & Cragg, G. M. Natural products as sources of new drugs over the nearly four decades from 01/1981 to 09/2019. *J. Nat. Prod.* **83**, 770–803 (2020).
- Fleming, A. On the antibacterial action of cultures of a penicillium, with special reference to their use in the isolation of *B. influenzae*. *Br. J. Exp. Pathol.* **10**, 226–236 (1929).
- Kennedy, J. et al. Modulation of polyketide synthase activity by accessory proteins during lovastatin biosynthesis. *Science* **284**, 1368–1372 (1999).
- Cohen, D. J. et al. Cyclosporine: a new immunosuppressive agent for organ transplantation. *Ann. Intern. Med.* **101**, 667–682 (1984).
- Weber, T. et al. antiSMASH 3.0-a comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic Acids Res.* **43**, W237–243 (2015).
- Khaldi, N. et al. SMURF: genomic mapping of fungal secondary metabolite clusters. *Fungal Genet. Biol.* **47**, 736–741 (2010).
- Rutledge, P. J. & Challis, G. L. Discovery of microbial natural products by activation of silent biosynthetic gene clusters. *Nat. Rev. Microbiol.* **13**, 509–523 (2015).
- Bok, J. W. & Keller, N. P. LaeA, a regulator of secondary metabolism in *Aspergillus* spp. *Eukaryot. Cell* **3**, 527–535 (2004).
- Mao, X. M. et al. Epigenetic genome mining of an endophytic fungus leads to the pleiotropic biosynthesis of natural products. *Angew. Chem.* **54**, 7592–7596 (2015).
- Xu, F. et al. A genetics-free method for high-throughput discovery of cryptic microbial metabolites. *Nat. Chem. Biol.* **15**, 161–168 (2019).
- Roux, I. et al. CRISPR-mediated activation of biosynthetic gene clusters for bioactive molecule discovery in filamentous fungi. *ACS Synth. Biol.* **9**, 1843–1854 (2020).
- Wolff, P. B. et al. Acurin A, a novel hybrid compound, biosynthesized by individually translated PKS-and NRPS-encoding genes in *Aspergillus aculeatus*. *Fungal Genet. Biol.* **139**, 103378 (2020).
- Chiang, Y.-M. et al. An efficient system for heterologous expression of secondary metabolite genes in *Aspergillus nidulans*. *J. Am. Chem. Soc.* **135**, 7720–7731 (2013).
- Tsunematsu, Y., Ishiuchi, K. I., Hotta, K. & Watanabe, K. Yeast-based genome mining, production and mechanistic studies of the biosynthesis of fungal polyketide and peptide natural products. *Nat. Prod. Rep.* **30**, 1139–1149 (2013).
- Yaegashi, J., Oakley, B. R. & Wang, C. C. Recent advances in genome mining of secondary metabolite biosynthetic gene clusters and the development of heterologous expression systems in *Aspergillus nidulans*. *J. Ind. Microbiol. Biotechnol.* **41**, 433–442 (2014).
- Chu, J., Vila-Farres, X. & Brady, S. F. Bioactive synthetic-bioinformatic natural product cyclic peptides inspired by nonribosomal peptide synthetase gene clusters from the human microbiome. *J. Am. Chem. Soc.* **141**, 15737–15741 (2019).
- Clevenger, K. D. et al. A scalable platform to identify fungal secondary metabolites and their gene clusters. *Nat. Chem. Biol.* **13**, 895–901 (2017).
- Navarro-Muñoz, J. C. et al. A computational framework to explore large-scale biosynthetic diversity. *Nat. Chem. Biol.* **16**, 60–68 (2020).
- Harvey, C. J. et al. HEX: a heterologous expression platform for the discovery of fungal natural products. *Sci. Adv.* **4**, eaar5459 (2018).
- Bian, G. et al. A clade II-D fungal chimeric diterpene synthase from *Colletotrichum gloeosporioides* produces dolasta-1 (15), 8-diene. *Angew. Chem. Int. Ed.* **57**, 15887–15890 (2018).
- Bian, G., Deng, Z. & Liu, T. Strategies for terpenoid overproduction and new terpenoid discovery. *Curr. Opin. Biotechnol.* **48**, 234–241 (2017).
- He, H. et al. Discovery of the cryptic function of terpene cyclases as aromatic prenyltransferases. *Nat. Commun.* **11**, 3958 (2020).
- Bian, G. et al. Releasing the potential power of terpene synthases by a robust precursor supply platform. *Metab. Eng.* **42**, 1–8 (2017).
- Nagamine, S. et al. Ascomycete *Aspergillus oryzae* is an efficient expression host for production of basidiomycete terpenes by using genomic DNA sequences. *Appl. Environ. Microbiol.* **85**, e00409–00419 (2019).
- Qin, B. et al. An unusual chimeric diterpene synthase from *Emericella variecolor* and its functional conversion into a sesterterpene synthase by domain swapping. *Angew. Chem. Int. Ed.* <https://doi.org/10.1002/anie.201509263> (2015).
- Ye, Y. et al. Genome mining for sesterterpenes using bifunctional terpene synthases reveals a unified intermediate of di/sesterterpenes. *J. Am. Chem. Soc.* **137**, 11846–11853 (2015).

28. Alberti, F. et al. Heterologous expression reveals the biosynthesis of the antibiotic pleuromutilin and generates bioactive semi-synthetic derivatives. *Nat. Commun.* **8**, 1831 (2017).
29. Zhang, X. et al. Genome mining and comparative biosynthesis of meroterpenoids from two phylogenetically distinct fungi. *Angew. Chem. Int. Ed.* **57**, 8184–8188 (2018).
30. Liu, C. et al. Efficient reconstitution of Basidiomycota diterpene erinacine gene cluster in Ascomycota host *Aspergillus oryzae* based on genomic DNA sequences. *J. Am. Chem. Soc.* **141**, 15519–15523 (2019).
31. Katayama, T. et al. Development of a genome editing technique using the CRISPR/Cas9 system in the industrial filamentous fungus *Aspergillus oryzae*. *Biotechnol. Lett.* **38**, 637–642 (2016).
32. Nodvig, C. S. et al. Efficient oligo nucleotide mediated CRISPR–Cas9 gene editing in *Aspergilli*. *Fungal Genet. Biol.* **115**, 78–89 (2018).
33. Wang, G. Q. et al. Biosynthetic pathway for furanosteroid demethoxyviridin and identification of an unusual pregnane side-chain cleavage. *Nat. Commun.* **9**, 1838 (2018).
34. Jin, F. J., Maruyama, J.-i., Juvvadi, P. R., Arioka, M. & Kitamoto, K. Development of a novel quadruple auxotrophic host transformation system by *argB* gene disruption using *adeA* gene and exploiting adenine auxotrophy in *Aspergillus oryzae*. *FEMS Microbiol. Lett.* **239**, 79–85 (2004).
35. Siemon, T. et al. Semisynthesis of plant-derived englerin A enabled by microbe engineering of guaia-6, 10 (14)-diene as building block. *J. Am. Chem. Soc.* **142**, 2760–2765 (2020).
36. Hertzberg, R. P. & Pope, A. J. High-throughput screening: new technology for the 21st century. *Curr. Opin. Chem. Biol.* **4**, 445–451 (2000).
37. Shapland, E. B. et al. Low-cost, high-throughput sequencing of DNA assemblies using a highly multiplexed Nextera process. *ACS Synth. Biol.* **4**, 860–866 (2015).
38. Lauchli, R. et al. High-throughput screening for terpene-synthase-cyclization activity and directed evolution of a terpene synthase. *Angew. Chem. Int. Ed. Engl.* **52**, 5571–5574 (2013).
39. Liu, W. & Jiang, R. Combinatorial and high-throughput screening approaches for strain engineering. *Appl. Microbiol. Biotechnol.* **99**, 2093–2104 (2015).
40. Yuan, Y. et al. Sesquiterpenoids produced by combining two sesquiterpene cyclases with promiscuous myxobacterial CYP260B1. *Chem. Bio. Chem.* **20**, 677–682 (2019).
41. Bian, G. et al. Metabolic engineering-based rapid characterization of a sesquiterpene cyclase and the skeletons of fusariumdiene and fusagrasimineol from *Fusarium graminearum*. *Org. Lett.* **20**, 1626–1629 (2018).
42. Sun, X. et al. Genome mining in *Trichoderma viride* J1-030: discovery and identification of novel sesquiterpene synthase and its products. *Beilstein J. Org. Chem.* **15**, 2052–2058 (2019).
43. Gao, L. et al. Identification of novel sesterterpenes by genome mining of phytopathogenic fungi *Phoma* and *Colletotrichum* sp. *Tetrahedron Lett.* **59**, 1136–1139 (2018).
44. Kudo, F., Matsuura, Y., Hayashi, T., Fukushima, M. & Eguchi, T. Genome mining of the sordarin biosynthetic gene cluster from *Sordaria araneosa* Cain ATCC 36386: characterization of cycloaraneosene synthase and GDP-6-deoxyaltrose transferase. *J. Antibiot.* **69**, 541–548 (2016).
45. Linstrom, P. J. & Mallard, W. G. *NIST Chemistry Webbook, NIST Standard Reference Database Number 69* (National Institute of Standards and Technology, 2022).
46. Del Valle, D. M. et al. An inflammatory cytokine signature predicts COVID-19 severity and survival. *Nat. Med.* **26**, 1636–1643 (2020).
47. Hortelano, S. Molecular basis of the anti-inflammatory effects of terpenoids. *Inflamm. Allergy Drug. Targets* **8**, 28–39 (2009).
48. Renner, M. K., Jensen, P. R. & Fenical, W. Mangicols: structures and biosynthesis of a new class of sesterterpene polyols from a marine fungus of the genus *Fusarium*. *J. Org. Chem.* **65**, 4843–4852 (2000).
49. Zhu, F. et al. In vitro reconstitution of mevalonate pathway and targeted engineering of farnesene overproduction in *Escherichia coli*. *Biotechnol. Bioeng.* **111**, 1396–1405 (2014).
50. Chen, R. et al. Systematic mining of fungal chimeric terpene synthases using an efficient precursor-providing yeast chassis. *Proc. Natl. Acad. Sci. USA* **118**, e2023247118 (2021).
51. Xie, S., Shen, B., Zhang, C., Huang, X. & Zhang, Y. sgRNACas9: a software package for designing CRISPR sgRNA and evaluating potential off-target cleavage sites. *PLoS ONE* **9**, e100448 (2014).
52. Katayama, T. et al. Forced recycling of an AMA1-based genome-editing plasmid allows for efficient multiple gene deletion/integration in the industrial filamentous fungus *Aspergillus oryzae*. *Appl. Environ. Microbiol.* **85**, e01896–01818 (2019).

Acknowledgements

We thank H. Liu (Wuhan Institute of Biotechnology) for his technical support in setting up the automatic high-throughput biofoundry workflow. This work was financially supported by the National Key R&D Program of China (grant nos. 2018YFA0900400 and 2021YFC2102600), the National Natural Science Foundation of China (grant nos. 31670090, 31800032 and 32070063) and the Medical Science Advancement Program (Clinical Medicine) of Wuhan University (grant no. TFLC2018002).

Author contributions

T.L., G.B. and Z.D. designed the experiments. Y.Y., Z.M. and P.Y. refactored terpenoid BGCs. Y.Y., S.C. and W.D. performed in vitro and in vivo screening of anti-inflammatory activity. S.C., H.H. and H.C. constructed the AO chassis and increased the titre of mangicol J. R.C. and Y.C. characterize the structure of terpenoids. Y.X., S.C., S.F., G.B. and T.L. analysed the data. G.B. and T.L. wrote the paper.

Competing interests

T.L., Y.Y., S.C., G.B., P.Y., Z.M. and R.C. have filed four patent applications (Chinese Patent Application Nos. 202010936790.X, 202111306130.4, 202111306130.4 and 202111261216.X) on the basis of this contribution. The remaining authors declare no competing interests.

Additional information

Extended data are available for this paper at <https://doi.org/10.1038/s41929-022-00762-x>.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41929-022-00762-x>.

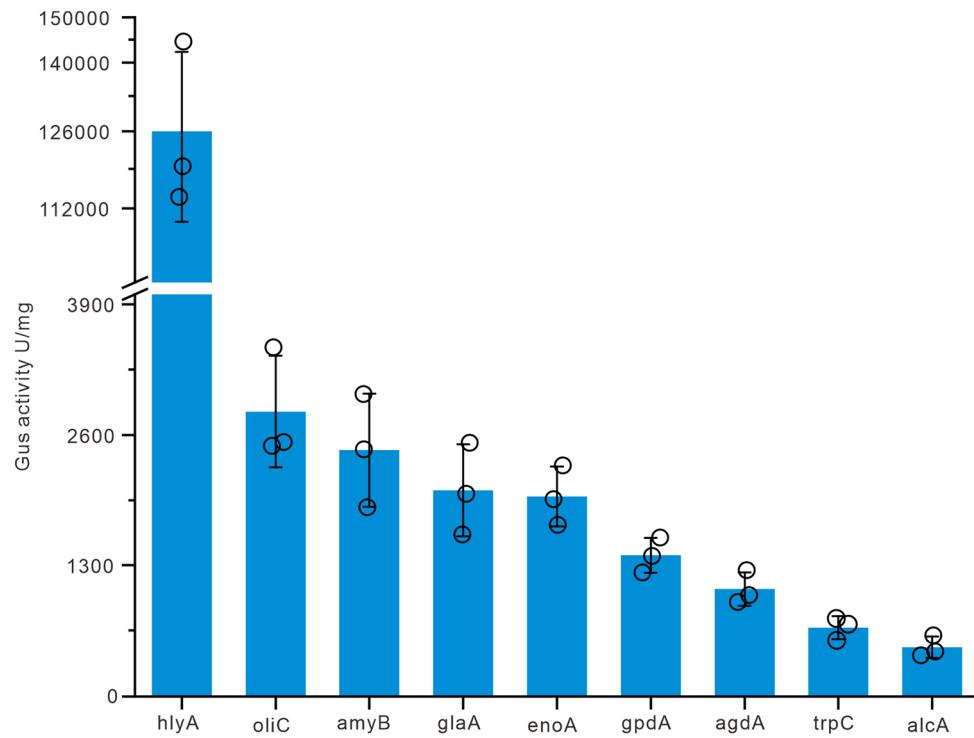
Correspondence and requests for materials should be addressed to Tiangang Liu.

Peer review information *Nature Catalysis* thanks Nancy Keller, Jose A Carrasco Lopez and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

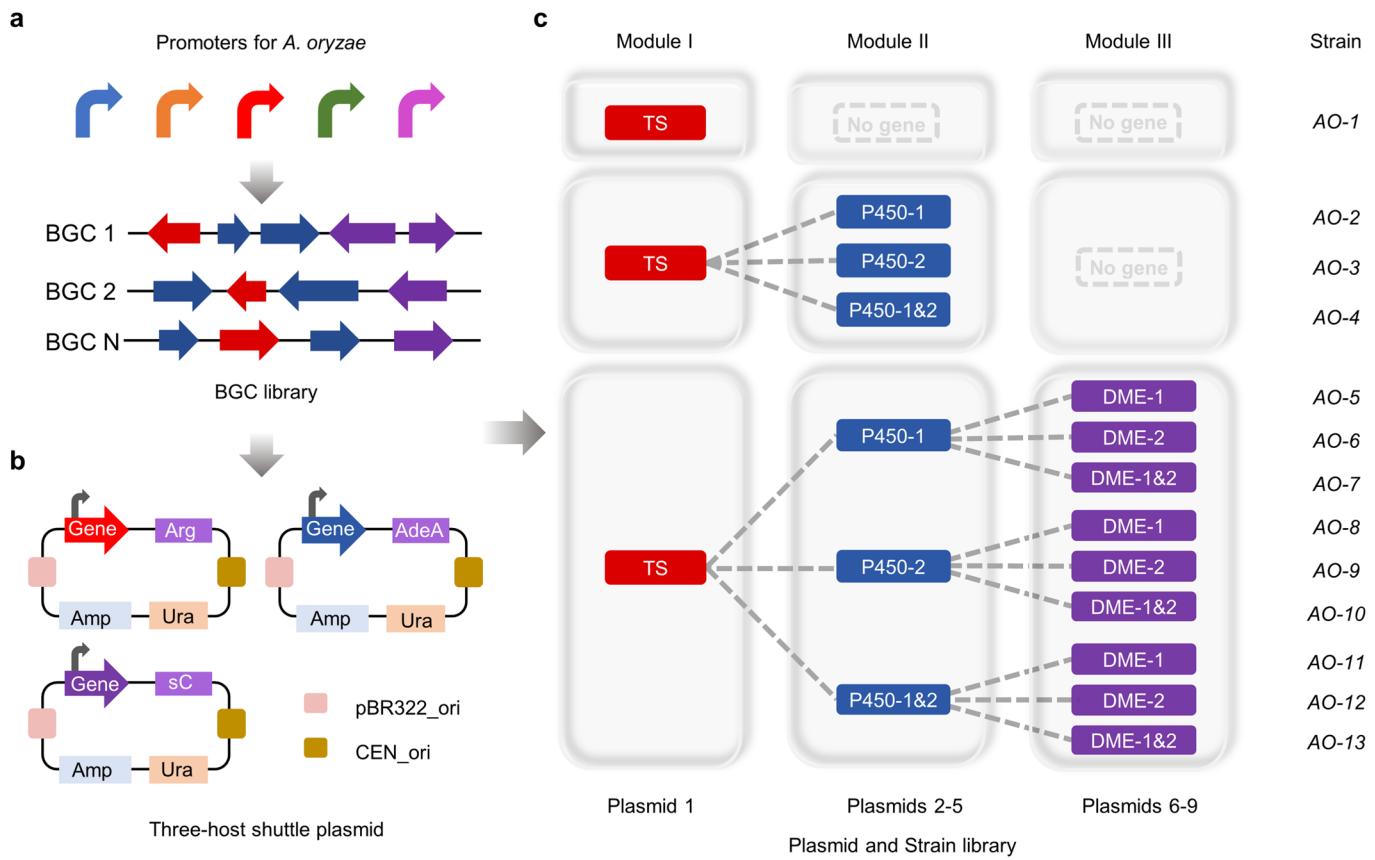
Reprints and permissions information is available at www.nature.com/reprints.

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

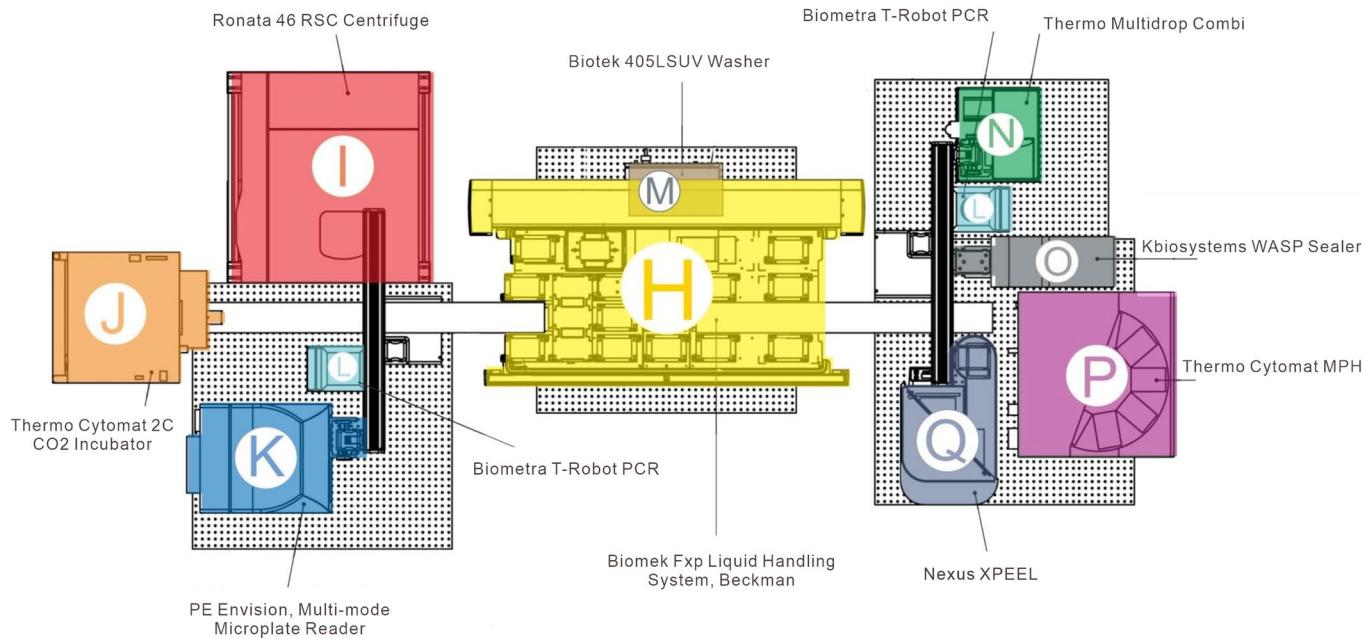
© The Author(s), under exclusive licence to Springer Nature Limited 2022



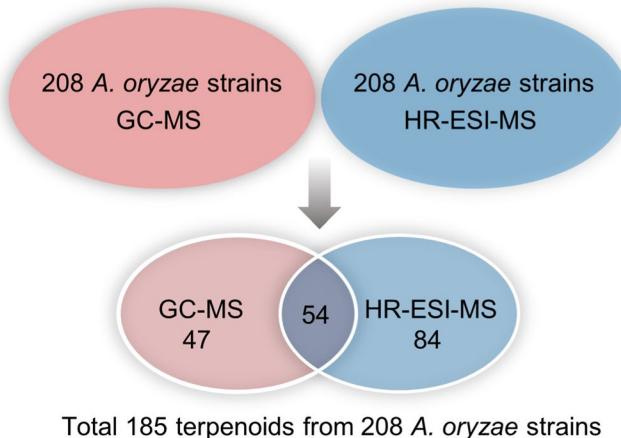
Extended Data Fig. 1 | Strength evaluation of selected promoters in AO-NSAR1. The strength of each promoter (black bars) was assessed based on GUS activity. All data are mean values of three independent experiments, and error bars indicate the standard deviation.



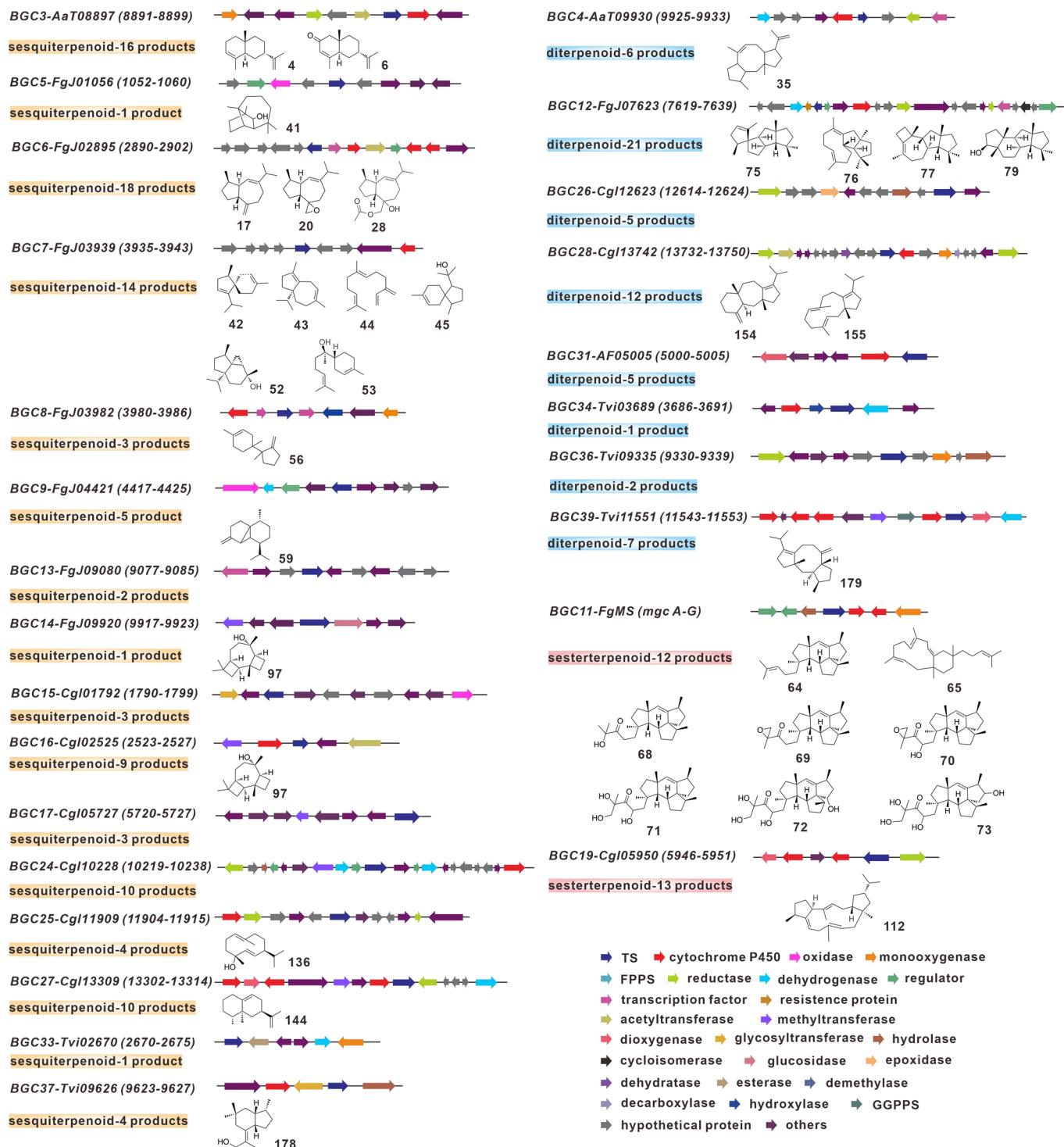
Extended Data Fig. 2 | Principles for plasmid and AO-strain construction. **a**, Strength-evaluated promoters and BGC library. **b**, Three-host shuttle plasmid library. **c**, Principle of plasmid and AO-strain construction and the strain library. Abbreviations: DME, Downstream modification enzyme.



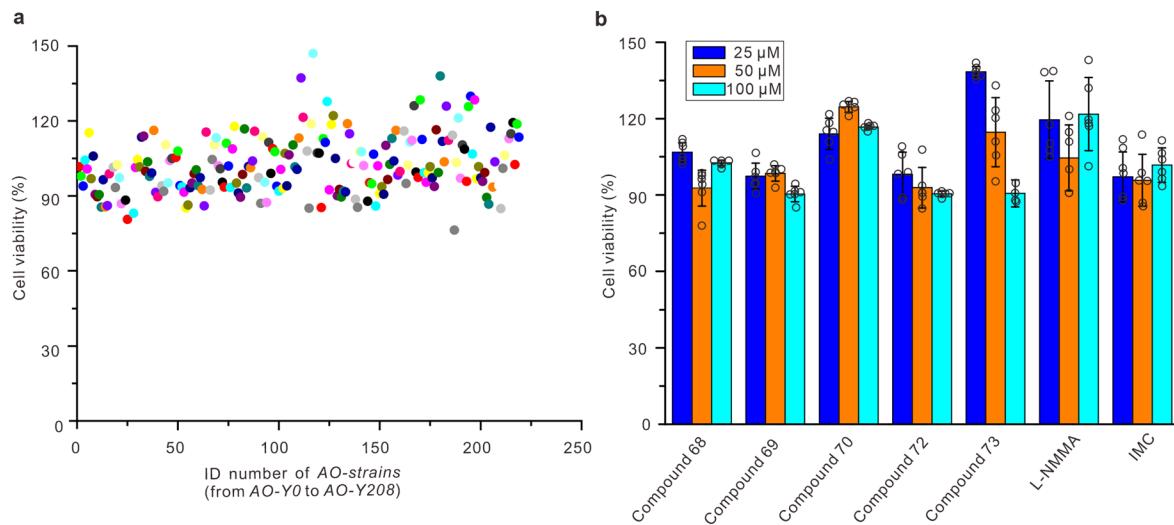
Extended Data Fig. 3 | Overview of Biomek FX® Laboratory Automation Workstation. The workstation is integrated with different functional devices as follows: 1) Biomek FX® Liquid Handling System in the core center (yellow, H), which is equipped with dual arms (a fixed AP96 module and a flexible 8 channel); 2) Biometra T-Robot PCR (white, L); 3) Thermo Multidrop Combi Dispenser (green, N); 4) Kbiosystems WASP Sealer (gray, O); 5) Thermo Cytomat MPH (purple, P); 6) Nexus XPEEL: automated plate seal removal system (dusty blue, Q); 7) Biometra 405L SUV Washer (transparent square, M); 8) Ronata 46 RSC Centrifuge (red, I); 9) Thermo Cytomat CO₂ Incubator (orange, J); 10) PE Envision, Multi-mode Microplate Reader (mazarine, K).



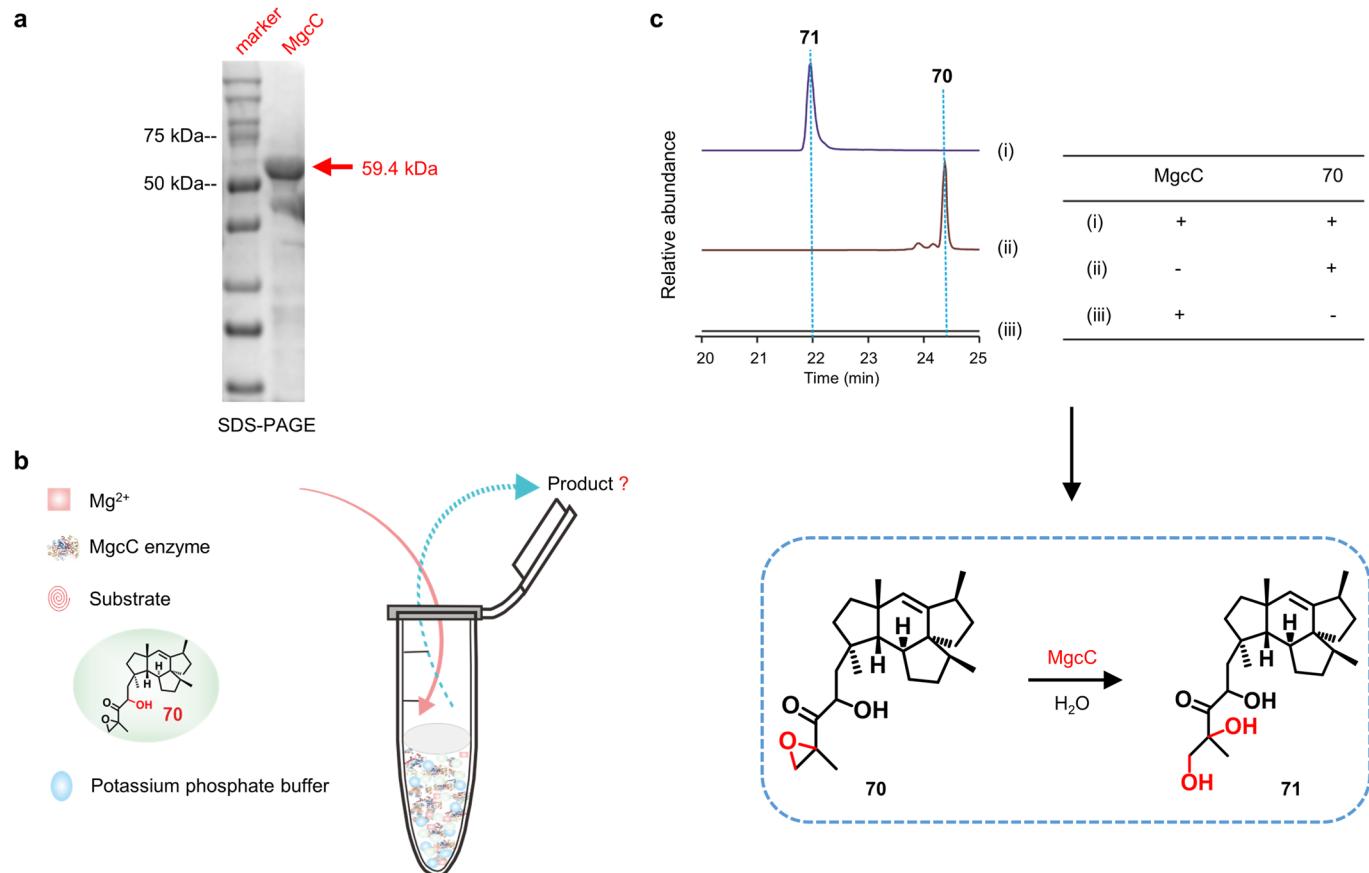
Extended Data Fig. 4 | Venn diagram of terpenoids detection number. The number terpenoids detected from *Aspergillus oryzae* strains using gas chromatography-mass spectrometry (GC-MS) and high-performance liquid chromatography (HPLC) coupled with high-resolution electrospray ionization mass spectrometry (HR-ESI-MS).



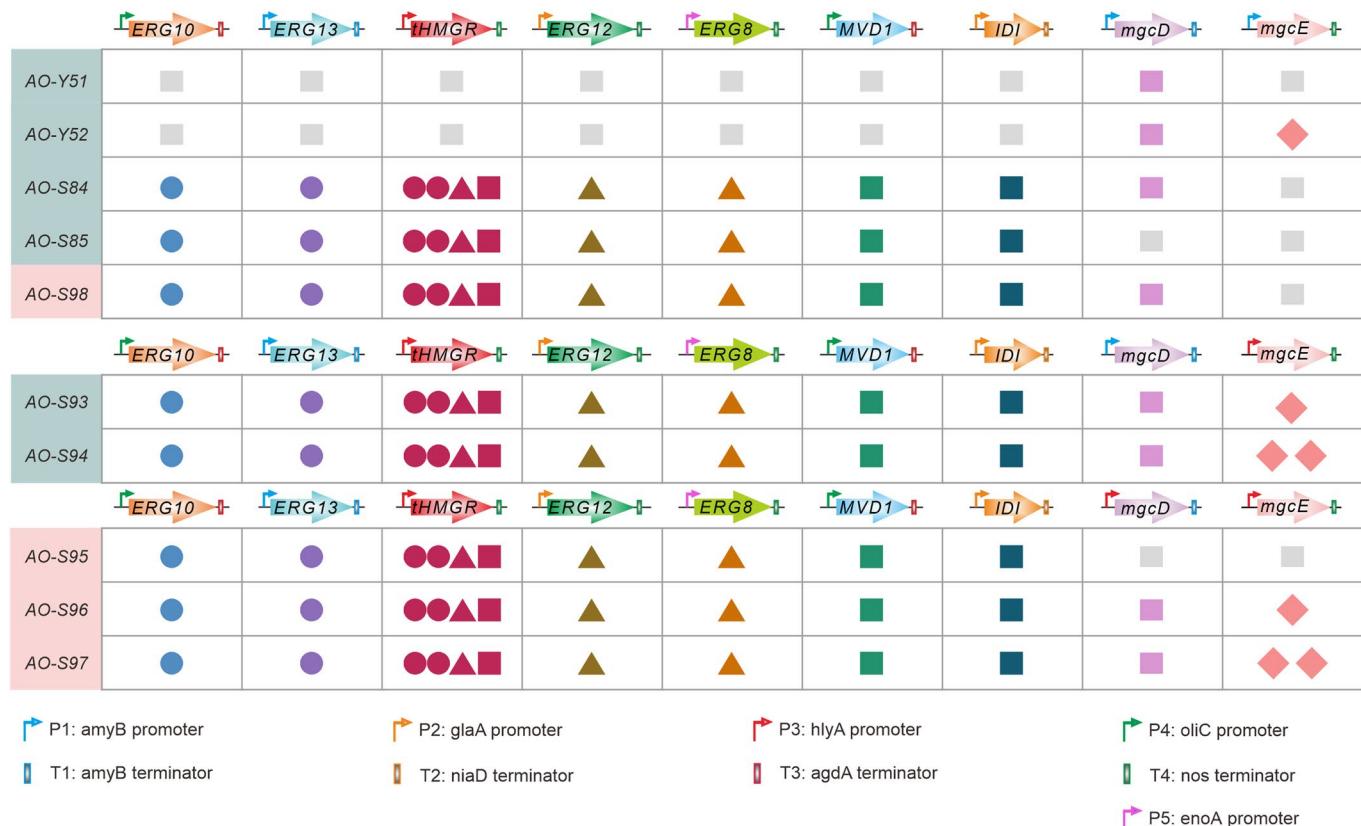
Extended Data Fig. 5 | Summary of terpenoids produced by terpenoid BGCs. 26 BGCs were detected the terpenoids, among which 16 sesquiterpenoid BGCs (yellow), 8 diterpenoid BGCs (blue) and 2 sesterterpenoid BGCs (red) were included, respectively. The number in, for example, “sesterterpenoid-12 products”, was represented the total products detected in each BGC.



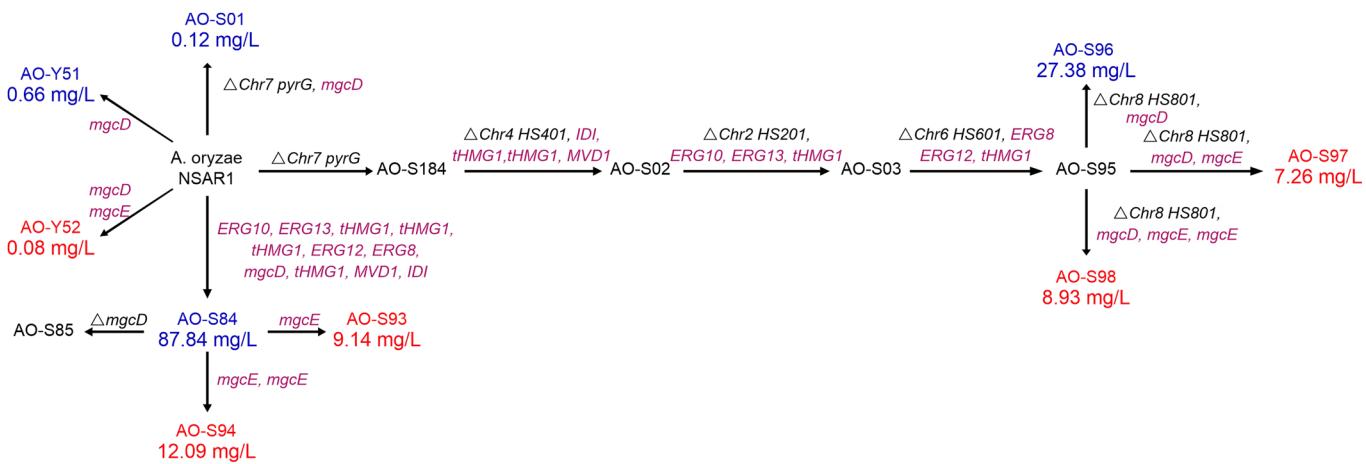
Extended Data Fig. 6 | Cell viability evaluation of compounds produced by AO-strains. **a**, Cell viability evaluation of RAW 264.7 microphage cells treated with crude products (final concentration, 500 μ g/mL) produced by AO-strains. **b**, Influence of different concentrations (25 μ M, 50 μ M, 100 μ M) of mangicols (68, 69, 70, 72 and 73) and L-NMMA and IMC towards the cell viability of RAW264.7 macrophage cells. L-NMMA, NG-Monomethyl-L-arginine, monoacetate salt (an inhibitor of nitric oxide synthase); IMC, indomethacin. Data (**a** and **b**) are presented as mean values \pm SD ($n=6$).



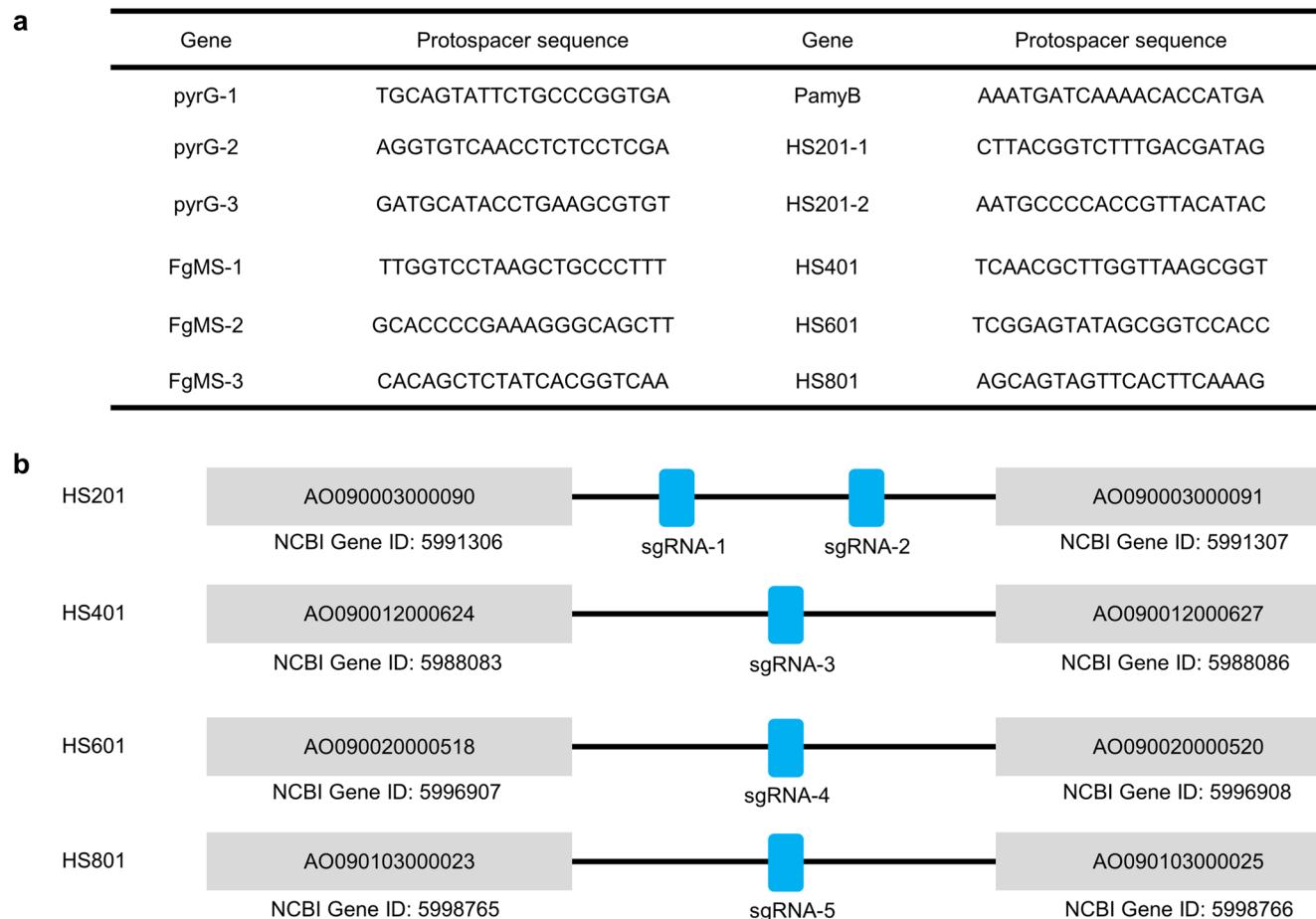
Extended Data Fig. 7 | Functional characterization of MgcC by *in vitro* assay. **a**, The expression verification of MgcC by SDS-PAGE. **b**, *In vitro* reaction system for functional verification of MgcC. The reaction included 0.5 μ M MgcC enzyme, 1 mM Mg²⁺, 200 μ M substrate (**70**) and 20 mM phosphate buffer (pH 7.4) was performed at 30 °C for 8 h. **c**, HR-ESI-MS detection of compound **71** produced by MgcC. Compound **71** were detected in the reaction contained MgcC and substrate **70** (i). The reactions, without MgcC (ii) or without substrate(ii), were unable to produce **71**. Three independent experiments were repeated with similar results.



Extended Data Fig. 8 | The strain for production of mangicdiene and mangicol J. Engineered strains via the endogenous MVA or overexpressed mevalonate pathways. Strains shaded in green were attained by gene randomly insertion, strains shaded in pink were attained by gene hotspot site integration. Circles, triangles, squares, and diamonds represent different plasmids, and different colors represent different genes.



Extended Data Fig. 9 | Flowchart of AO mutant construction in this study. Blue represents the production of mangicdiene, red represents the production of mangicol J, purple represents overexpressed genes.



Extended Data Fig. 10 | Protospacers used in this study. **a**, Protospacer sequences used in this study. **b**, High expression loci (hot spots, HS) and genetic information. The blue shaded area represents the sgRNA site. The gray shaded area represents the genes flanking the sgRNA site.

Corresponding author(s): Tiangang Liu

Last updated by author(s): Feb 16, 2022

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection GC-MS and HR-ESI-MS data collection software: Qual Browser Thermo Xcalibur 3.0.63 August 27, 2013; Flow cytometry data collection software: FCAP Array Software Version 3.0.1; NMR data collection software: Bruker, Version: AV III; Biomek software: 3.3.14.

Data analysis GC-MS and HR-ESI-MS data analysis software: Qual Browser Thermo Xcalibur 3.0.63 August 27, 2013; Flow cytometry data analysis software: FCAP Array Software Version 3.0.1; NMR data analysis software: MestReNova LITE.Ink. 5.3.1

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

Data supporting the findings of this study are available within the article and its Supplementary Information files. The accession numbers (NMDCN0000QH2 to NMDCN0000QH9, NMDCN0000QHA to NMDCN0000QHV, and NMDCN0000QJ0 to NMDCN0000QJ9) and nucleotide sequences for the characterized enzymes were deposited in the National Microbiology Data Center (<https://nmdc.cn/en>) will be available after publication. Relevant data has been provided through source data. Materials (e.g., plasmids, strains, compounds, etc.) in this study are available from the corresponding author upon reasonable request.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Anti-inflammatory activity assays in vitro and in vivo were determined with n=3 or n>3 as it is common practice in the field.
Data exclusions	No data were excluded from the analysis.
Replication	The in vivo anti-inflammatory assay were carried out with three to six replications, and the rest experiment were repeated with three parallel samples to make sure the data reproducibility.
Randomization	The strains, cells and mouses were randomly allocated into experimental groups.
Blinding	The investigators were blinded to group allocation during data analysis.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems		Methods	
n/a	Involved in the study	n/a	Involved in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies	<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input type="checkbox"/>	<input checked="" type="checkbox"/> Eukaryotic cell lines	<input type="checkbox"/>	<input checked="" type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology	<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging
<input type="checkbox"/>	<input checked="" type="checkbox"/> Animals and other organisms		
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants		
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data		
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern		

Antibodies

Antibodies used	First antibody: β-Actin, Mouse source, supplier name: abcam, catalog number: ab8226; First antibody: p-STAT3, Rabbit source, supplier name: CST, catalog number: 9145s; First antibody: STAT3, Rabbit source, supplier name: CST, catalog number: 12640s; Second antibody: hRP-Goat anti Rabbit, Goat source, supplier name: Elabscience, catalog number: E-AB-1003; Second antibody: hRP-Goat anti Mouse, Goat source, supplier name: Elabscience, catalog number: E-AB-1001.
Validation	Validation: First antibody β-Actin: https://www.abcam.cn/beta-actin-antibody-mabcam-8226-loading-control-ab8226.html . First antibody p-STAT3: https://www.cellsignal.com/products/primary-antibodies/stat3-d3z2g-rabbit-mab/12640?site-search-type=Products&N=4294956287&Ntt=12640s&fromPage=plp&_requestid=2993563 . Second antibody hRP-Goat anti Rabbit: https://www.elabscience.com/p-goat_antirabbit_igg_h_l_peroxidase_hrp_conjugated_-24504.html . Second antibody hRP-Goat anti Mouse: https://www.elabscience.com/p-goat_antimouse_igg_h_l_peroxidase_hrp_conjugated_-24572.html .

Eukaryotic cell lines

Policy information about cell lines	
Cell line source(s)	Cells (RAW 264.7 TIB-71™, Primary Umbilical Vein Endothelial Cells; Normal, Human (HUVEC) PCS-100-010™) used in this study were purchased from the American Type Culture Collection (ATCC, USA).
Authentication	The two cells line are both authenticated by morphology (microscopic observation after culture) and immunofluorescence

identification.

Mycoplasma contamination

All cell lines tested negative for mycoplasma contamination.

Commonly misidentified lines
(See [ICLAC](#) register)

No commonly misidentified cell lines was used in this study.

Animals and other organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research

Laboratory animals

Eight-week-old female ICR (CD-1®) mice (Harlan Sprague Dawley Inc.); Seven-week-old male C57BL/6J mice (Zhejiang Weitong Lihua Experimental Animal Technology Co., Ltd, Zhejiang, China).

Wild animals

This study did not involve wild animals.

Field-collected samples

The study did not involves samples collected from the field.

Ethics oversight

The animal care and use were adhered to the Chinese National Guidelines for Ethical Review of Animal Welfare. Animals were handled according to the Guidelines of the China Animal Welfare Legislation, and the study protocol was approved by the Institutional Animal Care and Use Committee (IACUC) of Renmin Hospital of Wuhan University Human Research Ethics Committee (IACUC issue no. 20201220A).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Flow Cytometry

Plots

Confirm that:

- The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- All plots are contour plots with outliers or pseudocolor plots.
- A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation

Describe the sample preparation, detailing the biological source of the cells and any tissue processing steps used.

Instrument

Identify the instrument used for data collection, specifying make and model number.

Software

Describe the software used to collect and analyze the flow cytometry data. For custom code that has been deposited into a community repository, provide accession details.

Cell population abundance

Describe the abundance of the relevant cell populations within post-sort fractions, providing details on the purity of the samples and how it was determined.

Gating strategy

Describe the gating strategy used for all relevant experiments, specifying the preliminary FSC/SSC gates of the starting cell population, indicating where boundaries between "positive" and "negative" staining cell populations are defined.

- Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.