Activation of an unconventional meroterpenoid gene cluster in *Neosartorya glabra* leads to the production of new berkeleyacetals

Tao Zhang\(^a,^\dagger\), Jun Wan\(^a,^\dagger\), Zhajun Zhan\(^b\), Jian Bai\(^a\), Bingyu Liu\(^a\), Youcai Hu\(^a,^*\)

\(^a\)State Key Laboratory of Bioactive Substance and Function of Natural Medicines, Institute of Materia Medica, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing 100050, China

\(^b\)College of Pharmaceutical Science, Zhejiang University of Technology, Hangzhou 310014, China

Received 9 September 2017; received in revised form 28 October 2017; accepted 11 November 2017

**KEY WORDS**

*Neosartorya glabra*; Meroterpenoid; Berkeleyacetals; Genome mining; Cryptic gene cluster; Biosynthesis

**Abstract**

Fungal genomes carry many gene clusters seemingly capable of natural products biosynthesis, yet most clusters remain cryptic or down-regulated. Genome mining revealed an unconventional paraherquonin-like meroterpenoid biosynthetic gene cluster in the chromosome of *Neosartorya glabra*. The cryptic or down-regulated pathway was activated by constitutive expression of pathway-specific regulator gene *berA* encoded within *ber* biosynthetic gene cluster. Chemical analysis of mutant *Ng*-OE: *berA* extracts enabled the isolation of four berkeleyacetal congeners, in which two of them are new. On the basis of careful bioinformatic analysis of the coding enzymes in the *ber* gene cluster, the biosynthetic pathway of berkeleyacetals was proposed. These results indicate that this approach would be valuable for discovery of novel natural products and will accelerate the exploitation of prodigious natural products in filamentous fungi.

© 2018 Chinese Pharmaceutical Association and Institute of Materia Medica, Chinese Academy of Medical Sciences. Production and hosting by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

---

\(^*\)Corresponding author.
E-mail address: huyoucai@imm.ac.cn (Youcai Hu).
\(^\dagger\)These authors made equal contribution to this work.

Peer review under responsibility of Institute of Materia Medica, Chinese Academy of Medical Sciences and Chinese Pharmaceutical Association.
1. Introduction

Meroterpenoids constitute an important family of hybrid natural products partially derived from terpenoid pathways, which have remarkably commercial and research values due to their diverse arrays of bioactivities and complex molecular architectures. 3,5-Dimethyl-lorsellinic acid (DMOA), an aromatic tetraketide intermediate widely synthesized by fungi, especially in the family Trichocomaceae, is further transformed into varied fascinating meroterpenoids.

Berkeleyacetals are heavily oxidized DMOA-derived meroterpenoids identified from *Penicillium rubrum* Stoll, and possess a unique and congested pentacyclic ring skeleton. Analogues, including minolutelides, berkeleydione, berkeleytrione, dhillorides, and paraherquonquin have been isolated from several fungi in the genus *Penicillium*. Interestingly, berkeleyacetals, berkeleydione and berkeleytrione reportedly inhibited matrix metalloproteinase-3 and caspase-1 effectively, in which berkeleyacetal C and berkeleydione were tested in antitumor screen for human cell line assay in NCI Developmental Therapeutics Program. Considering their biological properties, biosynthetic studies or engineering of natural products with the berkeleyacetal scaffold could definitely contribute toward clarifying the bioprocesses and facilitate the development of promising anticancer pharmaceuticals.

Genome mining of sequenced fungi has yielded new natural products with interesting bioactivities and structures, and a group of down-regulated or cryptic biosynthetic gene clusters were characterized and elucidated. Previous studies have demonstrated that genome mining, particularly the induction of down-regulated or cryptic secondary metabolic pathways by overexpressing of cluster-specific transcriptional activator genes, is a promising and more targeted strategy.

*Neosartorya glabra* was reported to produce numbers of natural products, such as glabamycins A–C, sartoryglabrin A–C, neosaphenols A and B, methoxyvermistatin, vermistatin, penicillide, purpactin, phialophoriol, etc. Whole genome sequencing of *N. glabra* revealed that its genome contains 17 polyketide synthetases (PKSs) gene clusters, 4 nonribosomal peptide synthetases (NRPSs) gene clusters, 3 PKS-NRPS hybrid gene clusters, and 5 terpene gene clusters, which far exceeds the total number of known natural products isolated from *N. glabra*. Interestingly, we mined the genome of *N. glabra* and identified a paraherquonquin-like meroterpenoid biosynthetic gene cluster that contains 5 additional genes other than prh cluster in *P. brasiliatinum* NBRC 6234 for paraherquonquin biosynthesis. Therefore, the unconventional gene cluster suggests it might be responsible for production of novel meroterpenoids. In this study, we demonstrated that overexpression of a pathway-specific regulator gene berA encoded within a down-regulated or cryptic, meroterpenoid biosynthetic gene cluster in *N. glabra* could stimulate the activation in a relatively straightforward fashion and 4 berkeleyacetal congeners or derivatives were characterized. This work provides the strategy to activate cryptic gene clusters by overexpressing pathway-specific regulator gene and help broaden our knowledge of the mechanism and pathway engineering of berkeleyacetals.

2. Materials and methods

2.1. Strains and cultivation conditions

The *N. glabra* strain was obtained from China General Microbiological Culture Collection Center and was used as the parental strain in our study. Both the wild-type and its mutant strains were grown on MEPA (3% malt extract broth, BD; 0.3% soy flour, 1.5% agar) for both secondary metabolites production and mRNA extraction at 28 °C. For gene overexpression in *N. glabra*, potato dextrose agar (BD) with 1.2 mol/L sorbitol and 400 μg/mL G418 was used for protoplast regeneration and antibiotic resistance selection. *Escherichia coli* Trans1-T1 was used for routine plasmid cloning. *Saccharomyces cerevisiae* strain BJ5464-NpgA (MATa ura3-52 his3-Δ200 leu2-Δ1 trp1 pep4-HIS3 prb1Δ1.6R can1 GAL) was used for in vivo yeast DNA recombination cloning and the yeast expression host. YPD (2% peptone, 1% yeast extract, 2% dextrose) was used for the routine growth of yeast strain BJ5464-NpgA and its derivatives at 30 °C. SD dropout medium was used for selection of plasmids transformed into *S. cerevisiae*. For protein expression under ADH2 promoter (ADH 2p) in *S. cerevisiae*, the yeast transformant was initially grown in the appropriate SD dropout liquid medium and then was transferred to the liquid YPD medium for further culture for 5 days. LB medium was used for culturing *E. coli*.

2.2. Sequencing and bioinformatic analysis

The genomic DNA of *N. glabra* used for sequencing was prepared from mycelium grown in stationary liquid culture (3% malt extract broth, BD; 0.3% soy flour). The shotgun sequencing was performed at Beijing Genomics Institute (Shenzhen, China) with the Illumina Hiseq. 2000 system. The contigs that assembled and annotated by SOAP denovo 1.05 were formatted to BLAST database for basic local BLAST search. AntiSMASH platform was used for genome mining and bioinformatic analysis of secondary metabolites biosynthetic clusters. Gene predictions were performed using the FGENESH program (Softberry) and manually checked by comparing with homologous proteins in the GenBank database. Functional domains in the translated protein sequences were predicted using Conserved Domain Search (NCBI) or InterproScan (EBI).

2.3. *N. glabra* RNA preparation, cDNA preparation, and reverse transcription-PCR (RT-PCR)

Mycelia of *N. glabra* and mutant Ng-OE: berA were inoculated into MEPA medium, incubated at 28 °C for 5 days, and collected for lyophilization. The total RNAs from culture of the wild type strain and mutant were extracted using the protocols as described previously. The genomic DNA was further removed by RNase-free DNase I (Takara). RNA was purified by RNAclean purification kit (Tiangen). RNA integrity was confirmed by electrophoresis on TAE buffer (Tris-acetate-EDTA) agarose gel. The first-strand cDNA was synthesized from 500 ng of total RNA by EasyScript® reverse transcriptase (Transgen) with random primers and oligo-dT18 primer (Takara) as described by the manufacturer. The gene expression level was analyzed by PCR using the specific primers listed in Supplementary information Table S1 and cDNA template. For BerA expression, PCR was performed with Q5 high-fidelity DNA polymerase (New England Biolabs) in the presence of 50 ng of reverse transcribed RNA. Primers are listed in Supplementary information Table S1.

2.4. Plasmid construction

Primers are listed in Supplementary information Table S1. Yeast expression plasmid pYET containing TRP1 auxotrophic marker.
was used for construction of the heterologous expression plasmid by in vivo homologous recombination in yeast. For polyketide synthase BerP expression, primers pairs BerP-S1 for/rev, BerP-S2 for/rev, and BerP-S3 for/rev were used to amplify three DNA fragments of berP cDNA and were transformed into S. cerevisiae BJ5464-NpgA simultaneously with Nle UPtrI digested pYET to create the plasmid pZT1. Yeast competent cell preparation and transformation were performed with a Frozen-EZ Yeast Transformation II kit (Zymo Research) according to the manufacturer’s protocol. Yeast plasmids were prepared by a Yeast Plasmid Miniprep kit (Solarbio) and transformed into E. coli strain Trans1-T1 for propagation and sequencing.

For construction of overexpression cassettes of berA, the gene berA was amplified from N. glabra genomic DNA using primers listed in Supplementary information Table S1. The constitutive gPdA promoter from Aspergillus nidulans (glyceraldehydes-3-phosphate dehydrogenase promoter) and G418 resistance gene fragment were amplified in which plasmids pRF-HUE and pYWl42 act as DNA templates. The three DNA fragments (gPdA, G418, and berA DNA fragments) were ligated into the linearized vector pET30a, which was digested with Hind III and EcoR I. The plasmids in the correct transformant screened by colony PCR were sequenced and used as template to amplify the overexpression cassette. Before transformation, the PCR products of overexpression cassette was recovered by a gel extraction kit (Omega, Cat. No. D2500-02) according to the manufacturer’s protocol and dissolved in STC buffer (1.2 mol/L sorbitol, 10 mmol/L CaCl2, 10 mmol/L Tris HCl, pH 7.5).

2.5 Fungal transformation and gene overexpression in N. glabra

Polyethylene glycol-mediated transformation of N. glabra was performed essentially as described previously for A. nidulans except that the protoplasts were prepared with 3 mg/mL lysing enzymes (Sigma–Aldrich) and 2 mg/mL tayalase (Takara). Briefly, fresh spores of N. glabra were collected and then induced to young germination with concentration of 106 spores mL−1 for 24 h at 28 °C with 150 rpm agitation. Young mycelia were harvested, washed twice with osmotic medium (1.2 mol/L MgCl2, 10 mmol/L sodium phosphate [pH 5.8]), and resuspended in the enzyme cocktail solution at 30 °C overnight. After washing twice with STC buffer (1.2 mol/L sorbitol, 10 mmol/L CaCl2, 10 mmol/L Tris–HCl, [pH 7.5]), protoplasts were gently mixed with DNA and incubated for 50 min on ice. One milliliter of PEG 4000 solution (60% PEG 4000, 50 mmol/L CaCl2, 50 mmol/L Tris–HCl [pH 7.5]) was added to 100 μL of protoplast mixture, and the mixture was incubated for 20 min at ambient temperature and spread on the regeneration selection medium (PDA, 1.2 mol/L sorbitol, 400 μg/mL G418). After incubation at 28 °C for 4–5 days, the transformants were inoculated on fresh PDB selection medium with stationary incubation for about 4 days to confirm the genotype by diagnostic PCRs after preparation of the genomic DNA. The specific primers used are shown in Supplementary information Table S1.

2.6 Chemical reagents and chemical analyses

All solvents and chemicals used this study are of analytical grade (for extraction) or LC–MS grade (for LC–MS analysis). Cultures of N. glabra, or S. cerevisiae cells were extracted with ethyl acetate. After 12,000 rpm, 10 min centrifugation (Eppendorf AG, MinimSpin, Hamburg, Germany), the supernatant organic phase was dried (Labconco Corporation, Dry Evaporators, Concentrators & Cold Traps, MO, USA) and solubilized in acetonitrile for LC–MS analyses. All LC–MS analyses were performed on a Waters ACQUITY H-Class UPLC–MS with a PDA detector and a QDA mass detector (ACQUITY UPLC® BEH, 1.7 μm, 50 mm × 2.1 mm, C18 column) using positive and negative mode electrospray ionization with a linear gradient of 5%–99% ACN–H2O (v/v, 0.02% formic acid) for 8 min followed by 99% ACN–H2O (v/v, 0.02% formic acid) for 4 min with a flow rate of 0.4 mL/min. X-ray data were collected using a Rigaku MicroMax 002+ instrument. The optical rotations were measured on a Jasco P2000 polarimeter, UV spectra were detected by a Jaslo V650 spectrophotometer (JASCO, Corporation, Tokyo, Japan). IR spectra were experimented on a Nicolet 5700 spectrophotometer via FT-IR microscope (Thermo Electron Scientific Instruments Corp.). NMR spectra was recorded on a Bruker AVIIIHD 600 (Bruker Corp., Karlsruhe, Germany) in DMSO-d6 at 600 MHz for 1H NMR and 150 MHz for 13C NMR, respectively, with solvent peaks used as references. HR-ESI-MS was measured on an Agilent 1100 series (Agilent Technologies, Ltd., Santa Clara, CA, USA). Sephadex LH-20 (Pharmacia Biotech AB, Uppsala, Sweden) was used for the open column chromatography. The medium pressure liquid chromatography was performed on Combi Flash Rf 2151320193 (Teledyne Isco, Lincoln NE, USA) and equipped with a dual pump gradient system, a UV preparative detector monitoring at 254 and 210 nm, a fraction collector system and a RP-C18 column (Sepaplash, sw080, 20–45 μm, Santai Technologies, Jiangsu, China). The semi-preparative HPLC was performed on SSI series 1500 (CoMetro Technology Ltd., NJ, USA) equipped with a DAD detector and a phenyl-hexyl column (250 mm × 10 mm, 5 μm, Phenomenex luna, CA, USA).

2.7 Fermentation, extraction and purification of secondary metabolites

The large-scale fermentation material of mutant Ng-OE was cultivated on MEPA medium at 28 °C for 7 days (280 plates, 140 mm) before collected into a 10 L vessel, and ultrasonic extracted with 7 L EtOAc (each 2 h × 4 times). The organic layer was evaporated to give a crude residue (ca. 37.79 g), which was then dissolved with ACN and subsequently partitioned by petroleum ether to yield a PE fraction (Fr. 1, 48.16 min) and 25%–60% for 35 min, 60% for 15 min and 70% for 20 min) to yield ten subfractions (Fr. 2.1–Fr. 2.10). Fr. 2.8 (346.3 mg) was then chromatographed on sephadex LH-20, eluted with MeOH giving eight subfractions (Fr. 2.81–Fr. 2.88). Fr. 2.84 (226.7 mg) was subjected to a semi-preparative HPLC to yield compounds 2 (0.9 mg, tR = 36.86 min) and 3 (44.3 mg, tR = 18.15 min). Fr. 3 (1.0215 g) was subjected to a RP-18 CC eluted with an ACN–H2O (20:80, 50:50 and 100:0, v/v), and acetone, to give four fractions (Fr. 1–Fr. 4). Fr. 2 (2.9747 g) was applied to a RP-18 CC (eluted with an ACN–H2O, 25% for 5 min, 25%–60% for 35 min, 45% for 15 min and 70% for 20 min) to yield ten subfractions (Fr. 2.1–Fr. 2.10). Fr. 2.8 (346.3 mg) was then chromatographed on sephadex LH-20, eluted with MeOH giving eight subfractions (Fr. 2.81–Fr. 2.88). Fr. 2.84 (226.7 mg) was subjected to a semi-preparative HPLC to yield compounds 2 (0.9 mg, tR = 36.86 min) and 3 (44.3 mg, tR = 18.15 min). Fr. 3 (1.0215 g) was subjected to a RP-18 CC eluted with an ACN–H2O, 45% for 7 min, 45%–85% for 20 min, 85% for 25 min and 100% for 20 min) to yield seven subfractions (Fr. 3.1–Fr. 3.7). Fr. 3.4 (154.2 mg) was then chromatographed on a semi-preparative HPLC to yield compounds 1 (4.4 mg, tR = 108.16 min) and 4 (1.6 mg, tR = 132.17 min).

2.7.1 Berkeleyacetal D (1)

Light yellow amorphous powder; [α]D25 31 (c 1.66, MeCN; UV (MeCN) λmax (logε) nm 208 (3.99), 267 (3.84). IR νmax 3065, 2983, 2908, 1767, 1707, 1604, 1671, 1397, 1318, 1297, 1262, 1222, 1155, 1116, 981, 932, 871, 768, 709 cm−1; For 1H and 13C NMR spectra were experimented on a Nicolet 5700 spectrophotometer via FT-IR microscope (Thermo Electron Scientific Instruments Corp.). NMR spectra was recorded on a Bruker AVIIIHD 600 (Bruker Corp., Karlsruhe, Germany) in DMSO-d6 at 600 MHz for 1H NMR and 150 MHz for 13C NMR, respectively, with solvent peaks used as references. HR-ESI-MS was measured on an Agilent 1100 series (Agilent Technologies, Ltd., Santa Clara, CA, USA). Sephadex LH-20 (Pharmacia Biotech AB, Uppsala, Sweden) was used for the open column chromatography. The medium pressure liquid chromatography was performed on Combi Flash Rf 2151320193 (Teledyne Isco, Lincoln NE, USA) and equipped with a dual pump gradient system, a UV preparative detector monitoring at 254 and 210 nm, a fraction collector system and a RP-C18 column (Sepaplash, sw080, 20–45 μm, Santai Technologies, Jiangsu, China). The semi-preparative HPLC was performed on SSI series 1500 (CoMetro Technology Ltd., NJ, USA) equipped with a DAD detector and a phenyl-hexyl column (250 mm × 10 mm, 5 μm, Phenomenex luna, CA, USA).
spectroscopic data, see Table 1; HR-ESI-MS (positive-ion mode) m/z 449.1567 [M + Na]^+ (Calcd. for C_23H_26O_7Na, 449.1571).

2.7.2. 11-epi-Berkeleyacetal C (2)

Light yellow amorphous powder; [α]_D^25 = −28.8 (c 0.66, MeCN; UV (MeCN) \(λ_{max} \text{ (logε)} \) nm 200 (3.49), 227 (3.52), 270 (3.79), IR \(ν_{max} \) 3082, 2983, 2938, 1786, 1716, 1710, 1659, 1455, 1392, 1372, 1294, 1198, 1128, 1080, 1008, 931, 872, 845, 591, 539 cm\(^{-1}\); For \(^1\)H and \(^13\)C NMR spectroscopic data, see Table 1; HR-ESI-MS (positive-ion mode) m/z 465.1503 [M + Na]^+ (Calcd. for C_{24}H_{28}O_8Na, 465.1520).

2.8. X-ray crystal structure analysis

Colorless crystals of I were obtained in MeOH. Intensity data was collected at Rigaku MicroMax 002+ X-ray diffractometer equipped with a CCD, using CuKar radiation. The structures were solved by direct methods using SHELXLS-97. Refinements were performed with SHELXL-97 using full-matrix least-squares, with anisotropic displacement parameters used for all the non-hydrogen atoms. The H atoms were placed in the calculated positions and refined using a riding model. Molecular graphics were computed with PLATON. Crystallographic data (excluding structure factor tables) for the structure reported has been deposited with the Cambridge Crystallographic Data Center as supplementary publication No. CCDC 1567469 for I. Copies of the data can be obtained free of charge on application to CCDC, 12 Union Road, Cambridge CB1 EZ, UK [fax: Int. + 44 (0) (1223) 336 033; email: deposit@ccdc.cam.ac.uk].

2.9. Crystallographic data for berkeleyacetal D (1)

C_{24}H_{28}O_8Na, MW = 426.45, orthorhombic system, space group \(P2_1_2_1_2_1\), cell parameter \(Z = 4, a = 10.908 \text{ (5) Å, } b = 12.474 \text{ (7) Å, } c = 15.344 \text{ (7) Å}; \alpha = \beta = \gamma = 90^\circ, V = 2078.8 \text{ (18) Å}^3, T = 295 K, \mu (\text{Cu Ka}) = 0.824 \text{ mm}^{-1}, 649 \text{ reflections measured, 3863 independent reflections. The final } R_1 \text{ value was 0.0381. The final } wR2 (F^2) \text{ value was 0.0977 } ([\geq 2\sigma(I)], S = 1.040. Flack parameter = 0.07 (8).}

3. Results and discussion

3.1. Genome mining of paraherquonin-like gene cluster in N. glabra and bioinformatic analysis

The Illumina HiSeq, 2500 sequencing of N. glabra CGMCC 32286 generated a total of ~1152 million bases with an average sequencing read length of 125 bases. Assembly of the unpaired shotgun sequence reads resulted in 66 contigs, which consists of 35.16 million nonredundant bases. The draft genome of N. glabra was then annotated using SOAP denovo program\(^{23}\). Bioinformatic analysis using antiSMASH\(^{24}\) revealed the organism could encode 12 biosynthetic gene clusters that may contain a NR-PKS, in which one biosynthetic gene cluster on contig 11 exhibited 20% similarity to terretonin. For our genomics-driven discovery of natural products and their biosynthetic mechanisms, we focused on the gene cluster designated as ber that resembles to paraherquonin biosynthetic cluster prh in P. brasilianum NBRC 6234. DNA sequence analysis of a contiguous ~58 kb ber locus (Table I) revealed the presence of 20 putative open reading frames (berA~T) that might be responsible for meroterpenoid biosynthesis. berP putatively encodes an iterative nonreducing PKS and has a domain architecture of SAT-KS-MAT-PT-ACP-CM-TE as ascertained by in silico analysis. Amino acid sequence alignment indicated that BerP shares 55% identity to AdrD, a known fungal iterative PKS from Penicillium roqueforti involved in Andrastin A biosynthesis\(^{25}\), followed by MpaC\(^{26}\), PrhL\(^{2}\), AusA\(^{28}\), and Trt4\(^{2}\), members of fungal DMOA-derived meroterpenoid PKSs.

Investigation of the flanking regions of the PKS gene allowed the discovery of other genes coding typical enzymes for meroterpenoid biosynthesis (Table 1). Other than genes with corresponding or homologous open reading frames in prh gene cluster for paraherquonin biosynthesis in P. brasilianum NBRC 6234, additional genes including

| Table 1 | Genes required for berkeleyacetals biosynthesis in Neosartorya glabra. |
|---------|--------------------------------------------------|
| Gene | No. | Proposed function | Coverage/identity | Protein homologue, organism | Accession No. |
| berA | 747 | C6 transcription factor protein | 90/33 | Transcription factor, T. benhamiae | XP_003012811 |
| berB | 377 | Thioredoxin-like protein AED1 | 53/67 | Thioredoxin, P. subrubescens 132785 | OKO89901 |
| berC | 431 | Cytochrome P450 monoxygenase | 99/83 | Pbh, P. brasiliananum NBRC 6234 | BAV69303 |
| berD | 174 | NAD-dependent epimerase | 100/67 | PbhC, P. brasiliananum NBRC 6234 | BAV69304 |
| berE | 579 | MFS general substrate transporter | 89/47 | PbhG, L. palustris CBS 459.81 | OCK75213 |
| berF | 170 | NAD-dependent epimerase | 90/31 | PbhC, P. brasiliananum NBRC 6234 | BAV69304 |
| berG | 239 | Terpene cyclase | 94/49 | PbhH, P. brasiliananum NBRC 6234 | BAV69309 |
| berH | 170 | O-acetyltransferase | 85/34 | AusA, Q. calidoustus | CEL11256 |
| berI | 645 | Flavin-containing monoxygenase-like | 91/54 | PbhI, P. brasiliananum NBRC 6234 | BAV69311 |
| berJ | 358 | FAD-dependent hydroxylase | 69/60 | PbhF, P. brasiliananum NBRC 6234 | BAV69307 |
| berK | 257 | Short-chain dehydrogenase/reductase | 100/63 | PbhI, P. brasiliananum NBRC 6234 | BAV69310 |
| berL | 279 | Methyltransferase | 100/64 | PbhM, P. brasiliananum NBRC 6234 | BAV69304 |
| berM | 643 | Flavin-containing monoxygenase-like | 97/67 | PbhN, P. brasiliananum NBRC 6234 | BAV693012 |
| berN | 512 | Cytochrome P450 monoxygenase | 91/44 | AusG, A. flavus NRRRL3557 | XP_002384778 |
| berO | 309 | UbiA prenyltransferase | 93/53 | PbhE, P. brasiliananum NBRC 6234 | BAV693006 |
| berP | 2458 | NR-PKS | 99/51 | PbhL, P. brasiliananum NBRC 6234 | BAV693013 |
| berQ | 133 | RutC family protein (soromerase) | 78/70 | RutC family protein, P. fici W106-1 | XP_007841478 |
| berR | 290 | Pyranylan-CoA dioxygenase | 98/74 | PrhA, P. brasiliananum NBRC 6234 | BAV69302 |
| berS | 434 | Cytochrome P450, putative | 94/66 | PrhD, P. brasiliananum NBRC 6234 | BAV69305 |
| berT | 446 | Cytochrome P450, putative | 100/64 | AusG, A. nidulans FGSC A4 | XP_682517 |

Note: T, Trichophyton; L, Lepidopterella; P, Penillium or Pestalotiopsis; A, Aspergillus.
berB, berH, berQ, and berT, exist in ber gene cluster. Sequence alignments and phylogenetic analysis to other meroterpenoid biosynthetic pathways, we speculated that ber biosynthetic cluster would be responsible for production of novel DMOA-derived meroterpenoids.

### 3.2. gpdA Promoter-controlled overexpression of berA induce metabolite production

Detailed analyses of the gene berA revealed the deduced BerA was related to the C6 transcriptional factor of *Trichophyton equinum* CBS 127.97 (EGE06077). To prove the concept that constitutive overexpression of a regulatory gene can lead to activation or up-regulation of the ber gene cluster, we amplified the putative activator gene berA from genomic DNA and gpdA promoter from plasmid pRF-HUE\textsuperscript{15} and cloned them into G418 resistant vector pYWL42. Transformation of *N. glabra* with the plasmid harboring berA under the control of the constitutive glyceraldehyde-3-phosphate dehydrogenase promoter gpdA of *A. nidulans* resulted in several mutant strains that harboring the expression cassette, and we named one desired mutant strain, Ng-OE: berA. RT-PCR analysis showed that transcription level of ber biosynthetic cluster in mutant Ng-OE: berA is much higher than that in wild type strain (Supplementary Information Fig. S1), in which four genes including berM, berP, berR, and berT were selected for detection. As a result, we sought to monitor the production of new metabolites in the mutant. Notably, UPLC coupled diode array and mass detectors revealed that the induced strain produces more compounds than wild type strain (Supplementary information Fig. S2). To get sufficient amounts of compounds for full structure elucidation, a large-scale of cultivation of the overexpressing mutant strain of *N. glabra* was carried out, which led to the isolation of two new compounds berkeleyacetal D (1) and 11-epi-berkeleyacetal C (2), along with two known ones, berkeleyacetal C (3)\textsuperscript{1} and purpurogenolide C (4)\textsuperscript{2} (Fig. 1). Berkeleyacetal D (1) was obtained as colorless crystals with its molecular formula of C\textsubscript{24}H\textsubscript{26}O\textsubscript{7} confirmed by HR-ESI-MS ([M+Na\textsuperscript{+}] 449.1567, Calcd. 449.1571), indicating an unsaturation of 12 degrees. Its IR spectrum gave absorption bands at 3065, 1666, 1604, 841 cm\textsuperscript{-1}, and 1767, 1707 cm\textsuperscript{-1}, suggesting the presence of olefinic groups and two types of carbonyl groups, respectively. The \textsuperscript{1}H NMR data (Table 2) displayed characteristic resonances for five methylenes (one oxygenated), seven methines, and two types of carbonyl groups, respectively. The \textsuperscript{13}C NMR and DEPT spectroscopic data (Table 2) exhibited 24 carbon signals, including five methylenes, corresponding to the \textsuperscript{1}H NMR data, two methylenes (one oxygenated), seven methines (three oxygenated and two olefinic) and ten quaternary carbons (two ester carbonyl). Preliminary inspection of NMR spectra of 1 suggested that it was a high oxygenated and condensed compound with structure similar to paraherquonin\textsuperscript{3} except that one oxygenated quaternary carbon (δC 73.3) and one methyl of paraherquinin was replaced by a methine (δC 40.0) and an epoxy group in 1.

![Figure 1 Structures of compounds 1–4.](image-url)

### Table 2 \textsuperscript{1}H NMR (600 MHz) and \textsuperscript{13}C NMR (150 MHz) spectral data of compounds 1-2 (δ in ppm, J in Hz, DMSO-d\textsubscript{6}).

| No. | 1       | 2       |
|-----|---------|---------|
|     | δ\textsubscript{H} | δ\textsubscript{C} | δ\textsubscript{H} | δ\textsubscript{C} |
| 1   | 162.5   | 161.7   | 144.8   | 117.6  |
| 2   | 5.79, d (1.2) | 155.0   | 150.0  |
| 3   | 57.7    | 59.1    | 37.2    | 44.3   |
| 4   | 2.33, dd (12.3, 4.2) | 26.7    | 28.3   |
| 5   | 1.62, (overlap) | 1.54, dd (14.6, 12.3) | 1.66, t (14.3) |
| 6a  | 1.54, dd (14.6, 12.3) | 1.66, t (14.3) |
| 6b  | 46.1    | 44.9    | 46.1    | 44.9   |
| 7   | 177.3   | 176.1   | 177.3   | 176.1  |
| 8   | 9.44, qd (6.7, 2.4) | 76.3    |
| 9   | 149.0   | 207.2   | 105.5   | 48.1   |
| 10  | 47.1    | 47.3    | 47.1    |
| 11  | 10.5    | 201.7   | 128.2   | 127.1  |
| 12  | 132.5   | 145.7   |
| 15  | 82.7    | 83.2    | 25.2    | 26.2   |
| 16  | 1.44, s  | 28.2    | 1.44, s  |
| 17  | 25.5    | 28.2    | 1.69, s  |
| 18  | 25.2    | 26.2    |
| 19  | 1.25, s  | 12.5    | 1.17, s  |
| 20  | 1.28, d (6.7) | 17.7    |
| 21  | 17.6    | 15.4, d (7.4) | 2.86, d (14.0, 4.5) | 91.1 |
| 22  | 97.2    | 97.9    | 2.69, d (4.5) |
| 23  | 23.2, s  | 23.5    | 23.2, s  |
| 24a | 55.1    | 56.8    | 5.89, d (5.4) |
| 24b | 2.57, d (5.4) |
| 24c | 2.57, d (5.4) |

\*The assignments were based on DEPT, \textsuperscript{1}H–\textsuperscript{1}H COSY, HSQC, HMBC experiments.
of an unsaturated cycloheptanone and a three-membered epoxy group and Me-20 substituent.

The $^1$H–$^1$H COSY correlations of H-5/H2-6, along with the HMBC correlations from H-5 to C-4, C-11(δC 105.5), C-12, and C-19, from H2-6 (δH 2.15 and 1.66) to C-4, C-7 (δC 46.1), C-8 (δC 177.3), C-12, and C-21 (δC 40.0), constructed the ring C. $^1$H–$^1$H COSY correlations of H-9/H2-20, and H-21/H-22, along with HMBC correlations from H-21 (δH 3.02) to C-7, C-8, C-10 (δC 149.0), C-11, and C-23 (δC 23.2), from H-22 (δH 6.24) to C-8, C-9 (δC 76.3, d), C-10, C-11, C-12 and C-21, revealed a $\gamma$-lactone jointed with an unsaturated pyrane ring by C-21 and C-22, in which the dioxygenated C-22 was jointed with C-8 and C-9. The HMBC correlations from H2-23 (δH 1.31) to C-6 (δC 28.3), C-7, C-8, and C-21, and from H2-20 (δH 1.34) to C-9, and C-10 suggested the methyl groups substituted at C-7 and C-9, respectively. Given that the unsaturation degrees, the oxygenated methine C-13 (δC 89.9), and the HMBC correlations from H-9, H-21, and H-22 to C-10 and C-11 disclosed that F ring was an $\Delta^{10}$ unsaturated furan ring.

The relative configuration of 1 was determined by the ROESY spectrum. ROESY correlations (Fig. 3) of H-13/H2-19, H-22/H2-21/H-23, along with the specific optical rotation compared with berkeleyacetal C, disclosed the $\alpha$-orientations of H-13, H2-19, H-21, H-22, and H-23. However, the ROESY correlations of H2-24 to H-5, and H4-20 to H-22 could not be observed, which made it difficult to determine the relative configurations of 4,24-epoxy group and Me-20 substituent.

With the aim to confirm the absolute configuration of compound 1, we have attempted to obtain its crystals. Fortunately, we succeeded in getting crystals of 1 from MeOH and performed the single-crystal X-ray diffraction experiment (Fig. 4). The final refinement on CuKα data resulted in a Flack parameter of 0.07 (8) allowing unequivocal assignments of the absolute configuration of the chiral carbons to be 4R, 5R, 7R, 21S, 22R, 9R, 13S. As a result, the configuration of 4,24-epoxy was confirmed as $\beta$-orientation, which was different to most reported berkeleyacetal-like meroterpenoids, while the Me-20 was determined to be $\alpha$-oriented, which might be contributed to the formation of the F ring.

11-epi-Berkeleyacetal C (2) was obtained as light yellow amorphous powder that gave a molecular formula of C22H26O8, as deduced by HR-ESI-MS. The IR spectra disclosed the existence of trisubstituted olefinic groups (3082, 1659, and 845 cm$^{-1}$), and at least three carbonyl groups (1786, 1736, and 1710 cm$^{-1}$). The $^1$H NMR data (Table 2) revealed five methyls at δH 1.17, 1.31, 1.43, 1.69 (each 3H, s), and δH 1.34 (3H, d, $J = 7.4$ Hz), and two conjugated olefinic protons δH 6.43, and 6.10 (each 1H, d, $J = 1.6$ Hz). The $^{13}$C NMR and DEPT spectra featured 24 carbons including five methyls, two methylenes, seven methines (two oxygenated and two olefinic ones), and ten carbons (two carbonyl carbons, two ester ones and two oxygenated ones). The NMR data showed that it formed a similar structure of berkeleyacetal C (3)$^1$, except for the sharp chemical shifts of C-5 (Δδ 12.1 ppm), C-10 (Δδ 4.4 ppm), C-21 (Δδ 4.7 ppm). The $^1$H–$^1$H COSY correlations of H-5/H2-6, H-11/H-21/H-22, and H-9/H-20, along with the HMBC correlations from H-24 (δH 2.89 and 2.57) to C-3 (δC 150.0), C-4 (δC 59.1), and C-5 (δC 44.3), from H-19 (δH 1.17) to C-5, C-12 (δC 47.3), and C-13 (δC 201.7), from H-21 (δH 2.86) to C-7 (δC 44.9, s), C-8 (δC 17.1), C-10 (δC 207.2), C-11 (δC 48.1), and C-23 (δC 23.5), from H-22 (δH 6.29) to C-8, C-11, and C-21 (δC 41.1), and from H-20 (δH 1.34) to C-9 (δC 76.3), and C-10, determined that compound 2 composing the same planar structure of 3. Considering the chemical shifts of several positions, we deduced that the configuration of compound 2 differed from that of 3, which was verified by the ROESY correlations. The ROESY correlations (Fig. 3) of H-24a/H-5/H3-19/H-11, H-19/H-21/H-22/H-23, and the coupling constant of $J_{21/12} = 14.0$ Hz and $J_{21/22} = 4.5$ Hz revealed the trans-orientations of H-21 and H-11, and the cis-orientations of H-21 and H-22, indicating that H-11 was $\beta$-oriented. Therefore, compound 2 was an H-11 epimer of berkeleyacetal C.

Compounds 3 and 4 were identified by the comparisons of their spectroscopic data with those reported in the literatures.

3.3. Comparative analysis of gene cluster ber with paraherquonin biosynthetic cluster prh in P. brasilianum NBRC 6234

A more detailed bioinformatic analysis of the biosynthetic locus ber revealed adjacent genes that are highly homologous to previously reported prh cluster in P. brasilianum genome$^2$. Both the ber and prh clusters contain a predicted DMOA-biosynthesis encoding gene as well as other putative ORFs highly conserved across previously reported biosynthetic pathway for fungal meroterpenoids. Although the ber and prh biosynthetic clusters are rearranged and nonsyntenic, the majority of their shared gene products are $>40\%$ sequence identical, and the correspondence between each ber gene from N. glabra and the respective ORF from the prh biosynthetic locus of strain P. brasilianum NBRC
Figure 5 Schematic representation of the ber cluster and its nucleotide sequence comparison with the prh cluster from *Penicillium brasiliannum NBRC 6234*.

Figure 6 The full-length amino acid sequences of BerC, BerN, BerS, and BerT with other P450 oxygenases were used to construct a phylogenetic tree by the neighbor-joining method. The bootstrap scores are based on 1000 reiterations. The BerC and PrhB in paraherquonin biosynthetic pathway are shown as an outgroup. P450 oxygenases in berkeleyacetals and paraherquonin pathways are shown in red and blue color, respectively.

6234 is described (Table 1, Fig. 5). As shown in Fig. 5, compared to prh cluster, the ber locus is a larger biosynthetic cluster. Among them, BerB shares 67% and 39% amino acid identity to thioredoxin-like protein and AhpC antioxidant enzyme of *Penicillium subrubescens* CBS 132785 and *Pochonia chlamydosporia* 170, respectively. Also, BerB shares 27% amino acid identity to Pyr7 and G, which are responsible for the biosynthesis of meroterpenoids including pyripyropene A and anditomin. Amino acid sequence alignment indicated that BerT is homologous to AusG from *Aspergillus nidulans* FGSC A4 (XP_682517, 64% identity) and followed by the Trt6 from *Penicillium roqueforti* CECT 2905 (ART41207, 47% identity) to those proteins characterized in other reported PKS genes, including AndM in *A. stellatus* (54% identity to BerP), AusA in *A. nidulans* (53% identity to BerP), Trt4 in *A. terreus* (52% identity to BerP) and PrhD in *P. brasilianum* (Q12612, 47% identity)8,30,36. Also, the heteroreactivity of *berP* was performed to confirm that identified pathway in *N. glabra* was correctly annotated, in which 3-, 5-dimethylsarselic acid was isolated and characterized (Supplementary information Fig. 1). Moreover, due to the gene products of *ber* cluster including BerO, BerL, BerG, BerK, BerI, BerM, and BerR share high similarity (42%–64%) to those proteins characterized in and, aus, and prh biosynthetic clusters, which were shown to produce meroterpenoids including andito Fig, austinol and paraherquo41, and paraherquonin. We reasoned that preaustinoid A1 by BerK and BerI is supported by the high similarity sequence similarity to short-chain dehydrogenase and flavin-containing monoxygenase (57% protein similarity with AndC) and AndE. The cyclization and formation of tetracyclic protoaustinoid A (10) from epoxyfarnesyl-DMOA methyl ester (9) could be mediated by BerG, which has protein sequence similarity to discovered terpene cyclases (PbhH, 49% identity and AusL, 42% identity)30,36. The hypothetical formation of preaustinoid A by BerK and BerI is supported by the high similarity sequence similarity to short-chain dehydrogenase and flavin-containing monoxygenase (57% protein similarity with and homologues, AndC and AndE). Based on the Baeyer-Valliger oxidation functions of its homologues, we proposed that BerM undergoes Baeyer-Valliger oxygen insertion to generate α-lactone ring system in preaustinoid A (13). BLAST analysis showed that BerH has 76% and 74% amino acid identity to AusE and PrhA, which belong to multifunctional Fe (II/III)-ketoligase (α-KG)-dependent dioxygenases family42,43. We deduced that BerH encodes a dioxygenase that catalyzes the construction of cycloheptadiene moiety into berkeleydione (15) from preaustinoid A (13) via the same mechanism.
during paraherquonin biosynthesis. The multifunctional dioxygenase BerR is the central player in the consecutive oxidations and structural rearrangement from “7+6” bicycle skeleton in 13 to “6+7” bicycle skeleton in 15. Upon formation of the on-pathway tetracyclic intermediate berkeleydione (15), multistep oxidation at distinct carbon atoms of substrates are required to afford the final berkeleyacetal D (1) and 11-epi-berkeleyacetal C (2). Previous studies demonstrated that cyclopropane formation of natural products could be catalyzed by cytochrome P450 oxygenases, examples including fumagillin, cytochalasins, aureothin. Similarly, we propose that the oxidative modifications at exo-methylene position of berkeleydione (15), and C4–C24 epoxidation of berkeleyacetal B (19) are likely to be catalyzed by the two cytochrome P450 oxygenases, BerN and BerT. As shown in Fig. 1, chemical structure of berkeleyacetal D and paraherquonin are similar except the generation of C4–C24 epoxide in former compound, indicates the parallel hidden biosynthetic mechanisms. This implies that BerT is most likely dedicated to the reaction for there is no homologue of BerT in prh gene cluster. The berkeleyacetals were previously isolated from fungal strains P. rubrum Berkeley Pit, P. purpureogenum MHZ113, and a marine mangrove-derived Penicillium sp. MA-37 which should be P. verruculosum MA-37 with 99% ITS identity. However, it has not been reported from strain P. brasiliense to the best of our knowledge.

BerN and BerS exhibits 47% and 25% protein identity to BerT, respectively, and this implies BerN possibly could be responsible for the oxidation from exo-methylene on C-22 to an aldehyde in compound 16 (Fig. 7), the epoxide-containing off-pathway compound 22-epoxyberkeleydione (21) could be the shunt product generated simultaneously. The existence of 21 is supported by the isolation and characterization of 22-epoxyberkeleydione in strain Penicillium sp. MA-37 and P. minioluteum. The BerS is proposed to involve the C-7 oxidation and elimination of the methyl ester group followed by the spontaneous decarboxylation of the β-keto acid, which shares 34% amino acid identity to the well-known multifunctional P450 oxygenase Trt6. This coincides with the proposed role of Trt6 being involved in the formation of terretonin H during terretonin biosynthesis. The last P450 monoxygenase might be participating in ether bond formation in compound berkeleyacetal D (1) through dehydration. The hypothesis proposed is supported by the high sequence similarity of BerC to PrhB encoded by prh biosynthetic cluster (82% protein identity), and the structural similarity of the product berkeleyacetal D to paraherquonin. Also, there is no P450 oxygenase homologues in other meroterpenoids biosynthesis to the best of our knowledge. As shown in Fig. 7, two epimerases including BerD or BerF could be proposed involving the rearrangement of intermediates, and the similarity of BerD or BerF to the AusH (28% amino acid identity) in the austinol pathway supports the proposed function of two proteins. Therefore, the biosynthetic pathway from tetracyclic intermediate berkeleydione (15) to highly oxygenated berkeleyacetals is proposed: BerN, a P450 monoxygenase, catalyzed an oxygenation coupled to generation of aldehyde group, followed by structural rearrangement and a second epoxidation that results in the conversion of berkeleyacetal A to berkeleyacetal B.
The P450 oxygenases BerS and BerC are believed to involve the conversions of berkeleyacetel C and berkeleyacetel D, respectively. In addition, the various stereochemistry at C6 and C13 in berkeleyacetals may due to the enolization of C10 carbonyl and keto-enol tautomerization.

4. Conclusions

We have identified a cryptic or down-regulated meroterpenoid gene cluster ber by genome mining, and successfully developed a strategy to activate the gene cluster by overexpressing pathway-specific regulator gene in N. glabra. As a result, we were able to isolate and characterize four berkeleyacetal derivatives (1–4). Moreover, bioinformatic analysis of the ber gene cluster was performed which helped to uncover a number of intriguing aspects of berkeleyacetals biosynthetic pathway. Further investigation on the multifunctional P450 oxygenases, including gene disruption of berkeleyacetals biosynthetic pathway. Further investigation on the multifunctional P450 oxygenases, including gene disruption of berkeleyacetals biosynthetic pathway. Further investigation on the multifunctional P450 oxygenases, including gene disruption (e.g., ΔberT, ΔberN and ΔberC), on-pathway intermediates characterization and in vitro biochemistry, should be performed to conclusively solve the problem.

Acknowledgments

This work was supported financially by the National Natural Science Foundation of China (No. 81522043), CAMS Initiative for Innovative Medicine (2017-12M-4-004), and the Thousand Young Talents Program of China. Youcai Hu is the Union Scholar in PUMC. We are grateful to Prof. Wenbin Yin (Institute of Microbiology, Chinese Academy of Sciences, Beijing, China) sharing plasmid pWYL42, and the Department of Instrumental Analysis of Peking University Medical College for the spectroscopic measurements and Dr. Ningbo Gong for solving the crystal structure of I.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at https://doi.org/10.1016/j.apsb.2017.12.005.

References

1. Matsuda Y, Abe I. Biosynthesis of fungal meroterpenoids. Nat Prod Rep 2016;33:26–53.
2. Geris R, Simpson TJ. Meroterpenoids produced by fungi. Nat Prod Rep 2009;26:1063–94.
3. Stierle DB, Stierle AA, Patacini B. The berkeleyacetals, three meroterpenes from a deep water acid mine waste Penicillium. J Nat Prod 2007;70:1820–3.
4. Stierle A, Stierle D, Decato D. Redetermination and absolute configuration of berkeleydione. Acta Crystallogr E Crystallogr Commun 2015;71:o248.
5. Stierle DB, Stierle AA, Hobbs JD, Stokken J, Clardy J. Berkeleydione and berkeleydione, two new bioactive metabolites from an acid mine organism. Org Lett 2004;6:1049–52.
6. Centko RM, Williams DE, Patrick BO, Akhtar Y, Garcia Chavez MA, Wang YA, et al. Dihilirofides E–N, meroterpenoids produced in culture by the fungus Penicillium purpurogenum collected in Sri Lanka: structure elucidation, stable isotope feeding studies, and insecticidal activity. J Org Chem 2014;79:3327–35.
7. de Silva ED, Williams DE, Jayanetti DR, Centko RM, Patrick BO, Wijesundera RL, et al. Dihilirofides A–D, meroterpenoids produced in...
31. Matsuda Y, Awakawa T, Wakimoto T, Abe I. Uncovering the unusual D-ring construction in terretonin biosynthesis by collaboration of a multifunctional cytochrome P450 and a unique isomerase. *J Am Chem Soc* 2015;137:3393–401.

32. Frandsen RJ, Andersson JA, Kristensen MB, Giese H. Efficient four fragment cloning for the construction of vectors for targeted gene replacement in filamentous fungi. *BMC Mol Biol* 2008;9:70.

33. Sun J, Zhu ZX, Song YL, Dong D, Zheng J, Liu T, et al. Nitric oxide inhibitory meroterpenoids from the fungus *Penicillium purpurigenum* MHZ 111. *J Nat Prod* 2016;79:1415–22.

34. Itoh T, Tokunaga K, Matsuda Y, Fujii I, Abe I, Ebizuka Y, et al. Reconstitution of a fungal meroterpenoid biosynthesis reveals the involvement of a novel family of terpene cyclases. *Nat Chem* 2010;2:838–44.

35. Matsuda Y, Wakimoto T, Mori T, Awakawa T, Abe I. Complete biosynthetic pathway of anditomin: nature's sophisticated synthetic route to a complex fungal meroterpenoid. *J Am Chem Soc* 2014;136:15326–36.

36. Matsuda Y, Wakimoto T, Mori T, Awakawa T, Abe I. Unusual chemistries in fungal meroterpenoid biosynthesis. *Curr Opin Chem Biol* 2016;31:1–7.