

Brief Report

Establishment of recombineering genome editing system in *Paraburkholderia megapolitana* empowers activation of silent biosynthetic gene clusters

Wentao Zheng,[†] Xue Wang,[†] Haibo Zhou, Youming Zhang* Aiyong Li** and Xiaoying Bian*** 
Helmholtz International Lab for Anti-Infectives, Shandong University-Helmholtz Institute of Biotechnology, State Key Laboratory of Microbial Technology, Shandong University, Qingdao, Shandong 266237, China.

Summary

The Burkholderiales are an emerging source of bioactive natural products. Their genomes contain a large number of cryptic biosynthetic gene clusters (BGCs), indicating great potential for novel structures. However, the lack of genetic tools for the most of Burkholderiales strains restricts the mining of these cryptic BGCs. We previously discovered novel phage recombinases Red $\alpha\beta$ 7029 from Burkholderiales strain DSM 7029 that could help in efficiently editing several Burkholderiales genomes and established the recombineering genome editing system in Burkholderiales species. Herein, we report the application of this phage recombinase system in another species *Paraburkholderia megapolitana* DSM 23488, resulting in activation of two silent non-ribosomal peptide synthetase/polyketide synthase BGCs. A novel class of lipopeptide, haereomegapolitanin, was

identified through spectroscopic characterization. Haereomegapolitanin A represents an unusual threonine-tagged lipopeptide which is longer than the predicted NRPS assembly line. This recombineering-mediated genome editing system shows great potential for genetic manipulation of more Burkholderiales species to activate silent BGCs for bioactive metabolites discovery.

Introduction

With the advancement of next-generation sequencing methods, a growing number of bacterial genomic sequences become accessible. The whole genome sequencing data suggest that the bacterial genomes contain more biosynthetic gene clusters (BGCs) than the known natural products, which indicates that a large number of microbial BGCs are in 'silent' and reveals that there is still a huge potential for excavation of microbial natural product diversity (Cimermancic, *et al.*, 2014; Rutledge and Challis, 2015; Zarins-Tutt, *et al.*, 2016; Mukherjee, *et al.*, 2017). The current genome mining strategies are divided into two types according to the host: homologous expression in the original bacteria and heterologous expression in the heterologous host (Bachmann, *et al.*, 2014; Huo, *et al.*, 2019). The advantage of activation silent gene clusters in the original bacteria is that the original bacterium contains prerequisite enzymes and substrates for biosynthesis of final products. Several strategies including manipulation of biosynthetic genes and regulatory factors, promoter engineering, addition of potential stimuli and comparative metabolic analyses of cultures under various growth conditions were harnessed to trigger silent BGCs in the original bacteria. Nevertheless, the homologous expression still relies on the genetic manipulation of native strains to connect the BGCs and products.

The Burkholderiales belong to an order of β -proteobacteria, including highly diverse bacteria that exist in various ecological niches such as human lung, fungal mycelia, plant, water and soil. The Burkholderiales strains, especially genera *Burkholderia* and *Paraburkholderia*, produce diverse bioactive natural products, for example anticancer

Received 30 September, 2019; revised 12 December, 2019; accepted 30 December, 2019.

*For correspondence. E-mail zhangyouming@sdu.edu.cn (Y.Z.); Tel. +86-532-67721918; Fax +86-532-58631501. **E-mail ayli@sdu.edu.cn (A.L.); Tel. +86-532-67722918; Fax +86-532-58631501. ***E-mail bianxiaoying@sdu.edu.cn (X.B.); Tel. +86-532-67720928; Fax +86-532-58631501.

[†]These authors are contributed equally to this work.

Microbial Biotechnology (2020) 13(2), 397–405
doi:10.1111/1751-7915.13535

Funding information

This work was supported by the National Key R&D Programme of China (2017YFD0201400, 2019YFA0905700), National Natural Science Foundation of China (31670098, 31670097), the Shandong Provincial Natural Science Foundation, China (ZR2019JQ11, ZR2017BC059), the China Postdoctoral Science Foundation (2019M652373), the Recruitment Programme of Global Experts (Y.Z.); the Qilu Youth Scholar Startup Funding of SDU (to X.B.).

© 2020 The Authors. *Microbial Biotechnology* published by John Wiley & Sons Ltd and Society for Applied Microbiology.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

agent FK228, analogs burkholdacs/thailandepsin (Biggins, *et al.*, 2011; Wang, *et al.*, 2011; Liu, *et al.*, 2018) and antifungal pesticide pyrrolnitrin (el-Banna and Winkelmann, 1998); thus, Burkholderiales become an emerging source of bioactive molecules. Genome analysis demonstrates that a large amount of BGCs in Burkholderiales are still cryptic or silent, and biosynthesized products by these cryptic BGCs also stand for a major uncharted source for novel bioactive products (Liu and Cheng, 2014; Depoorter, *et al.*, 2016; Esmaeel, *et al.*, 2016; Esmaeel, *et al.*, 2018). Many cryptic BGCs belong to polyketide synthases (PKS) and non-ribosomal peptide synthetases (NRPS) mostly encoding small molecule natural products with potential biological activities (Hertweck, 2009; Sussmuth and Mainz, 2017). However, lack of efficient genetic tools in the non-model Burkholderiales species impedes deeply exploration of novel bioactive products.

Red/ET recombineering is a homologous recombination technique firstly used in *Escherichia coli* by using short homology arms (~50 bp) based on either λ phage Red α /Red β or the equivalent Rac prophage RecE/RecT recombinases (Fu, *et al.*, 2012), which is suitable especially for precise genome engineering such as promoter insertion that fuels genome mining. *Burkholderia thailandensis* and *Burkholderia pseudomallei* are naturally competent for DNA transformation, which allows marked deletion by using PCR products with long homology arms (~1000 bp) (Thongdee, *et al.*, 2008). The mutated λ Red proteins, Red α (R137Q), Red β (C118R) and Red γ (added 18 bp 5'-UTR) (Nakayama and Ohara, 2005) from *E. coli* are capable to modify naturally transformable *Burkholderia* species directly with short homology arms (40-45bp), but failed in other species (Kang, *et al.*, 2011). Recently, we discovered a pair of novel phage recombinases Red $\alpha\beta$ 7029 from Burkholderiales strain DSM 7029 (recently reclassified as *Schlegelella brevitalea* (Tang, *et al.*, 2019)), which can be used for efficient genome engineering in several non-naturally transformable Burkholderiales strains that lack effective genetic tools. With this recombineering system, it is convenient to perform exact *in situ* insertion of functional promoters in genomes through one-step recombination for cryptic BGCs mining and leads to the discovery of two novel classes of lipopeptides (Wang, *et al.*, 2018). Moreover, a recombineering system with a lambda Red-like operon (BAS) from *Pseudomonas aeruginosa* phage Ab31 was established to perform genome engineering of four *Pseudomonas* species efficiently (Yin, *et al.*, 2019). These recombineering-mediated genome editing systems are not only doable in the native bacteria but also effective in other close strains, thus provide a feasible alternative for the Burkholderiales strains which lack of native phage recombinase or recombineering system.

Paraburkholderia megapolitana A3 (= DSM 23488 = LMG 23650) was isolated from a moss *Aulacomnium palustre* sampled in north-eastern Germany (Mecklenburg-Pommern). It exhibited antifungal activity against several plant pathogens and plant-growth-promoting properties (Vandamme, *et al.*, 2007). The genome sequence is accessible in NCBI, and the antiSMASH analysis suggested 15 BGCs included; most of them show low similarity to the known BGCs (MIBiG) (Medema, *et al.*, 2015), but no natural product has been isolated from this strain (Eberl and Vandamme, 2016), indicating that DSM 23488 has great potential for the discovery of novel compounds. Using a potent exogenous promoter to drive core biosynthetic genes within a silent gene cluster could heighten the transcriptional level of the core genes and avoid touching the complex native regulatory networks, is a straightforward method to mine silent BGCs. Herein, we firstly establish the optimal genome editing system for *P. megapolitana* DSM 23488 by screening different recombineering systems from Burkholderiales and *Pseudomonas* species, and implement recombineering-assisted promoter insertion to awaken two silent BGCs. A class of threonine-tagged lipopeptide, haereomegapolitanin, is identified, and haereomegapolitanin A is found to have biosurfactant activity and promotion of bacterial cell motility.

Results and discussion

Sequence and analysis of P. megapolitana genome

The whole genome shotgun sequence of *P. megapolitana* DSM 23488 in NCBI (NZ_FOQU00000000.1) contains a total of 32 contigs in 32 scaffolds with the size of 7.6 Mbp. The genome contigs are gapped, could not provide complete BGCs; thus, we resequenced the whole genome using PacBio Sequel platform. Its genome is 7 627 463 bp in length with an average GC content 62.1 % for two chromosomes and a plasmid. The entire genome contains 6936 coding genes, with an average length of 1089 bp. There are 69 RNA genes identified in the genome, which are further partitioned into 57 transfer RNA (tRNA) genes, four 5S rRNA genes, four 16S rRNA genes and four 23S rRNA genes. The antiSMASH analysis of reassembled *P. megapolitana* genome shows nine BGCs in chromosome 1, six BGCs in chromosome 2 and no BGC in plasmid (Table S1) (Blin, *et al.*, 2017). Five NRPS and two NRPS-PKS hybrid clusters are highly different from the known BGCs deposited in MIBiG database (Medema, *et al.*, 2015), implying the novel structure biosynthetic capacity of DSM 23488. The genome features of DSM 23488 are summarized in Table 1. The genome sequence data have been submitted to the GenBank under accession number CP041743-CP041745.

Establishment of recombineering genome editing system in *P. megapolitana*

Bioinformatics analysis reveals neither Red $\alpha\beta$ -like nor RecET-like recombinase pair presents in the genome of *P. megapolitana* DSM 23488, which urges us to consider employing foreign phage recombinases. Our previous study showed Red $\alpha\beta$ 7029 from Burkholderiales strain DSM 7029 could efficiently mediate genome editing of two *Paraburkholderia* species (Wang, *et al.*, 2018). Considering *Pseudomonas* is closely related to Burkholderiales, the BAS system from *Pseudomonas* was also chosen for functional verification in DSM 23488. Thus, Red $\gamma\beta\alpha$ (*E. coli*), BAS (*Pseudomonas*) and Red $\alpha\beta$ 7029 (Burkholderiales) were employed to perform genetic modification in DSM 23488. Different combinations of recombinases were constructed on expression plasmids with broad host range origin pBBR1 and under the control of inducible *rhaR-rhaS* P_{Rha} promoter (Wang, *et al.*, 2018).

The functional characterization of five recombinase combinations was carried out by plasmid and genome modification assays in DSM 23488. In the plasmid modification assay (Fig. 1A), standard linear plus circular homologous recombination (LCHR) assays using linear dsDNA substrate (apramycin resistance gene, Apra) flanked by various length homology arms to replace the promoter region and recombinase genes on the recombinase expression construct. All of the five recombinase combinations are able to perform plasmid modification when the length of homology arms is longer than 75 bp, but with varying efficiencies. Red γ -Red $\alpha\beta$ 7029 functions the most efficiently in DSM 23488 with colony number per millilitre (cnpm) of ~ 160 (Fig. 1B), followed by Red γ -BAS with cnpm of ~ 100 (Fig. 1B). Eight recombinants of each assay were verified by restriction analysis, and all of them were correct.

In the genome modification assay, an apramycin antibiotic resistance gene flanked with homology arms of different length was used to substitute ~ 17 kb *hrrmC*

gene (2754164–2771137) on chromosome (Fig. 2A). Red γ -Red $\alpha\beta$ 7029 and Red γ -BAS both could mediate genome modification in DSM 23488 when the homology arms reach a length of 100 bp, with cnpm of ~ 140 and 60, respectively (Fig. 2B). Eight independent recombinants were evaluated as correct mutants by colony PCR analysis. Red $\gamma\beta\alpha$ recombinase combination derived from *E. coli* failed to function in DSM 23488 genome modification assay. The addition of Red γ into all recombinase combinations remarkably enhanced the recombination efficiency in DSM 23488, as shown in our previous investigations in other strains (Yin, *et al.*, 2015; Wang, *et al.*, 2018). The recombination efficiency increased when the length of homology arms increased. Red γ -Red $\alpha\beta$ 7029 is the most efficient recombinase combination of the tested five combinations in DSM 23488 and selected for genome manipulation to mine the silent BGCs.

Activation of silent BGCs in *P. megapolitana*

The cryptic NRPS/PKS BGCs were explored by comparative metabolite analysis between inactivated mutants, promoter-inserted activated mutants and DSM 23488 wild type (Table S3). The inactivated mutant of target BGC was constructed by replacement of the core biosynthetic gene, for example the NRPS/PKS gene by an antibiotic selection marker with its own promoter. In the construction of activated mutant, the same antibiotic marker together with its own constitutive promoter was precisely inserted upstream of the first biosynthetic gene to drive the whole BGC (Fig. 3 and Figs. S1-S2). Using this strategy, we generated inactivated and activated mutants of seven NRPS/PKS BGCs of *P. megapolitana* DSM 23488, respectively. Metabolite analysis of the seven inactivated mutants and the DSM 23488 wild type showed no difference, indicating that all these BGCs remain in silent state under the laboratory fermentation conditions. The possible reason could be lack of environmental pressures. In the promoter insertion activated mutants, products of BGC5 (m/z 958, 984, and 982 [M + H]⁺) and BGC9 (m/z 765 (1) and 545 (2) [M + H]⁺) were discovered by comparing metabolite profiles with inactivated mutants and the wild-type strain (Fig. 3). The recombineering-assisted promoter insertion triggered two silent BGCs and their respective products, proving this strategy again is efficient in genome mining of Burkholderiales species (Wang, *et al.*, 2018). But the products of the rest five BGCs are still undetected, the possible reason could be they are still in a very low abundance escaped metabolic detection despite promoter insertion increased expression level of biosynthetic genes.

Table 1. Genome features of *P. megapolitana* DSM 23488.

Characteristics	Number
Genome size (bp)	7 627 463
Genes	7005
CDS	6936
rRNA	12
tRNA	57
G + C (%)	62.1
Gene average length (bp)	1089
Gene clusters for secondary metabolites	15
NRPS	5
NRPS-PKS	2

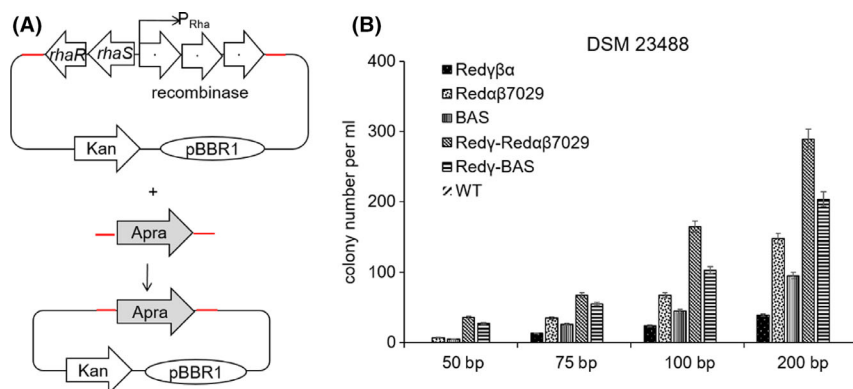


Fig. 1. The recombination efficiency of the plasmid modification assay mediated by different combinations of recombinases in *P. megapolitana* DSM 23488.

A. Diagram of plasmid modification assay (linear plus circular homologous recombination, LCHR) in DSM 23488.

B. Recombination efficiency comparison of five combinations of recombinases in DSM 23488. Red $\gamma\beta\alpha$ from *E. coli*; Red $\alpha\beta7029$ from *Burkholderia* strain DSM 7029; Red γ -Red $\alpha\beta7029$: Red γ from *E. coli* combined with Red $\alpha\beta7029$; BAS from *P. aeruginosa* phage vB_PaeP_Tr60_Ab31; Red γ -BAS: Red γ from *E. coli* combined with BAS. Error bars, SD; $n = 3$.

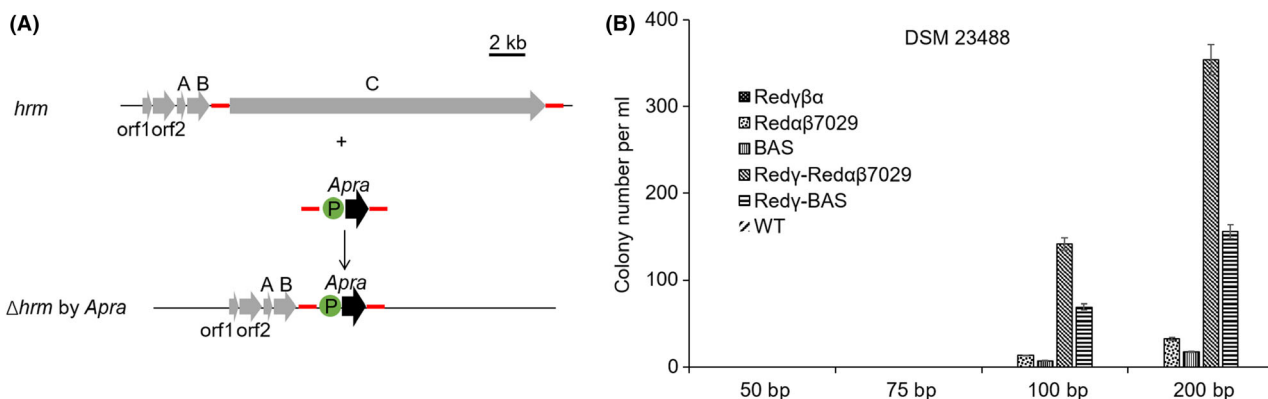


Fig. 2. The recombination efficiency of the genome modification assay mediated by different combinations of recombinases in *P. megapolitana* DSM 23488. A. Diagram of genome modification in DSM 23488. B. Recombination efficiency comparison of genome modification mediated by different recombinases. Error bars, SD; $n = 3$.

Structures and biosynthesis of lipopeptide haereomegapolitanins from *P. megapolitana*

The BGC9 shows high similarity to haereogladin BGC from *Burkholderia gladioli* pv. *agaricola* and haereoglumin BGC from *Burkholderia glumae* (Thongkongkaew, *et al.*, 2018). Thus, we name BGC9 as haereomegapolitanin (*hrm*) BGC. The *hrmC* encodes an NRPS containing five modules with a starter condensation (C_s)

domain, expecting to form a 5-amino acid lipopeptide FA-Thr-Ser-Leu-Ile-X (fatty acid, FA), which is discriminating to haereogladin and haereoglumin (Fig. 3). The deletion of *hrmC* in the BGC9 activated mutant leads to disappearance of BGC9 products, indicating that *hrmC* is the core gene of BGC9 (Fig. S2). The *hrmA* and *hrmB* encode an aminodeoxychorismate/anthranilate synthase and an aminodeoxychorismate synthase that are commonly used in para-aminobenzoic acid (PABA)

Fig. 3. Mining of the cryptic gene clusters in *P. megapolitana* DSM 23488.

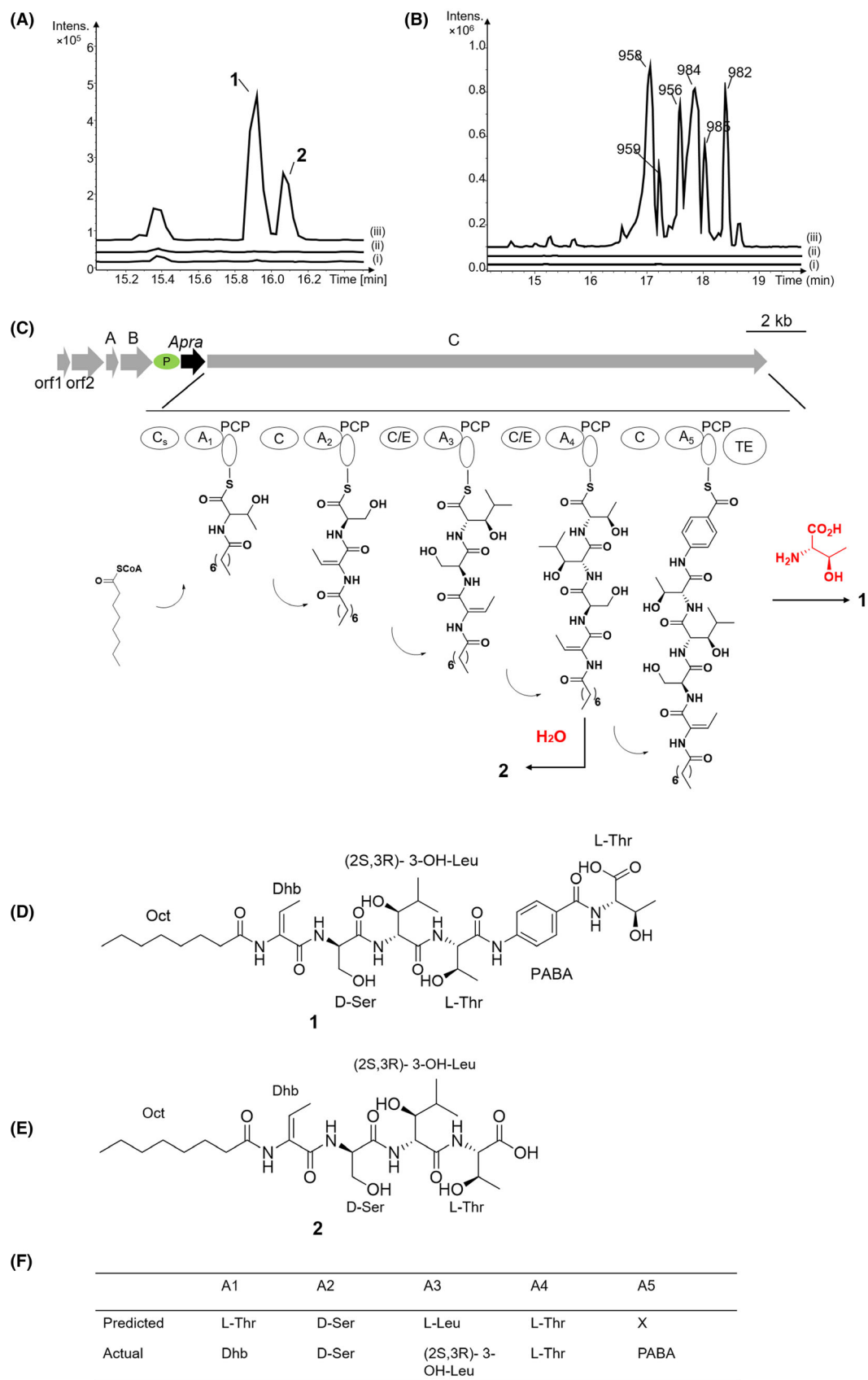
A. HPLC-HRMS analysis (BPC + 765 and 545) of extracts from (i) DSM 23488 wild type, (ii) BGC9 inactivated mutant of DSM 23488-DSM 23488:: Δ BGC9 and (iii) BGC9 activated mutant of DSM 23488-DSM 23488::P_{Apra}-BGC9; the m/z of **1** and **2** is 765 and 545, respectively.

B. HPLC-HRMS analysis of extracts from (i) DSM 23488 wild type, (ii) DSM 23488- Δ BGC5 and (iii) DSM 23488::P_{Apra}-BGC5.

C. The NRPS module architecture and model of biosynthetic pathway to **1** and **2**. BGC9 (*hrm*) encoding the haereomegapolitanins with promoter insertion (P-*Apra*), and predicted NRPS module and A domain for *hrm*. A, adenylation domain; C, condensation domain; C/E, dual condensation/epimerization domain; C_s , starter condensation domain; PCP, peptidyl carrier domain.

D. and E. Complete structures of haereomegapolitanin A (**1**) and haereomegapolitanin B (**2**).

F. The predicted and actual substrates of each module are listed. Dhb, 2,3-dehydrobutyric acid; PABA, para-aminobenzoic acid.



biosynthesis, which is supported by the predicted PABA substrate binding specificity of A5 domain (the fifth adenylation domain). The deletion of *hrmA* or *hrmB* in BGC9 activated mutants reduced the yield of the haereomegapolitanins (Fig. S3). Two products, haereomegapolitanin A (**1**) and B (**2**), were detected in the promoter knock-in mutant, and **1** was isolated as brown amorphous powder and analysed by HRMS and NMR for structure elucidation (Table S5). The molecular formula of **1** was determined to be $C_{36}H_{56}N_6O_{12}$ according to the HRESIMS ion at m/z 765.3986 $[M + H]^+$ (calculated 765.4029) (Fig. S2D–E). The ^{13}C and DEPT spectra showed signals of 7 carbonyl, 1 olefinic methine, 4 aromatic methine, 3 quaternary olefinic carbons, 8 methine, 7 methylene and 6 methyl protons (Fig. S4 and S5). Analysis of 1H - 1H COSY correlation indicated a spin system of octanoate (Oct), 2,3-dehydrobutyric acid (Dhb), serine (Ser), 3-OH-leucine (3-OH-Leu), two units of threonine (Thr) and PABA (Fig. S6–S8). Based on the HMBC correlations between amide proton and adjacent carbonyl, the peptide sequence was clarified as Oct - Dhb - Ser - 3-OH-Leu - Thr - PABA - Thr (Fig. S9). These data were further confirmed by mass spectrometry (Fig. S10A). The absolute configurations of amino acids were determined by Marfey's method as D-Ser, L-Thr and (2R,3S)-3-OH-Leu, which is in agreement with bioinformatics analysis of the C domains (Fig. 3, Fig. S11). Compound **2** was not purified enough for NMR analysis due to the low yield. Its HR ESIMS ion at m/z 545.3136 $[M + H]^+$ suggested molecular formula as $C_{25}H_{44}N_4O_9$ (calculated 545.3181). The comparative analysis of HRMS/MS spectra of **1** and **2** indicated the absence of two C-terminal monomers PABA and Thr in **2** (Fig. S10B). Haereomegapolitanins A and B are composed of six and four amino acid residues, respectively, differing from the five adenylation domains in BGC9. Haereomegapolitanin A contains an additional C-terminal Thr, which is rare but recently observed in some lipopeptides, haeroglandin and haereoglumin, from *Burkholderia* species (Thongkongkaew, *et al.*, 2018). The BGC9 is similar to haeroglandin BGC; thus, the mechanism of the terminal Thr and chain termination may be identical. The Thr tag might be introduced by a nucleophilic attack using its amino group to the thioester bond in PCP5 domain, forming peptide linkage between the Thr amino group and PABA (Fig. 3C). But the exact mechanism of chain termination by addition of an amino acid instead of hydrolysis is still an obscure. Haereomegapolitanin B lacks a PABA compared to predicted peptide backbone. The thioester bond in the PCP4 domain is cleaved by hydrolysis, resulting in early release of peptide chain to form the final product **2** (Fig. 3C). The insufficient supplement of PABA, the substrate of A5 domain, could be partial reason for the early release of peptide chain at PCP4 domain.

Biological activity of haereomegapolitanins

The lipopeptides with unusual threonine tags are conserved in mushroom and rice pathogenic *Burkholderia*, which could promote bacterial infection in the host. These biosurfactant-functional lipopeptides could achieve host colonization by mediating bacterial swarming and biofilm formation in host infections (Thongkongkaew, *et al.*, 2018). We performed a swarming assay with DSM 23488 wild-type strain and the BGC9 activated mutant strain (P_{AprA}-BGC9) to verify the possible effects of this novel lipopeptides on bacterial cell motility. This assay was performed on the soft Potato Dextrose Agar (PDA) (0.5% agar) plates; the BGC9 activated mutant strain swarmed out and colonized a bigger area compared to the wild-type strain and the inactivated strain mutant, which indicates that the BGC9 activated mutant shows swarming ability (Fig. 4A and B). This result indicates that haereomegapolitanins promote the swarming ability of the BGC9 activated mutant.

We likewise found that the BGC9 activated mutant strain also produced lower amount of biofilm compare to wild-type or inactivated mutant (Fig. 4C). Biosurfactants are known to have important regulatory effects on biofilms formation (Raaijmakers *et al.*, 2010). Then, we tested the surfactant potential of haereomegapolitanins by a toluene emulsion assay. Only the haereomegapolitanin A formed the emulsion (Fig. 4D), which indicates that the surfactant activity is related to haereomegapolitanin A. In addition, the result of contact angle shows that haereomegapolitanin A has surfactant properties, and haereomegapolitanin B does not display tensioactive activity. The critical micelle concentrations (CMC), a characteristic of lipopeptides, were determined from contact angle graph, and the CMC value for haereomegapolitanin A is 1.25 mM (Fig. 4E). Based on these results, we speculate that the haereomegapolitanin A is biosurfactant which is required for certain stages of invading host. Unexpectedly, haereomegapolitanin A showed neither *in vitro* antifungal activity against several fungi (*Rhizoctonia solani*, *Aspergillus fumigatus*, *Aspergillus oryzae*, *Aspergillus flavus* and *Candida albicans*) nor *in vitro* inhibition activity against several bacteria (*Staphylococcus aureus*, *Bacillus subtilis*, *Escherichia coli* and *Pseudomonas aeruginosa*). Haereomegapolitanin A only plays a role in the process of invading host and does not possess antifungal activity. The antifungal activity of DSM 23488 in the primitive living environment needs bioactive compounds that could be biosynthesized by other cryptic BGCs.

Colonization of bacteria in eukaryotic hosts is a common condition in many ecosystems and is a prerequisite for symbiotic and pathogenic relationships in various organisms. Although these processes are important in the living environment of bacteria and other organisms, little information about their internal mechanisms and chemical

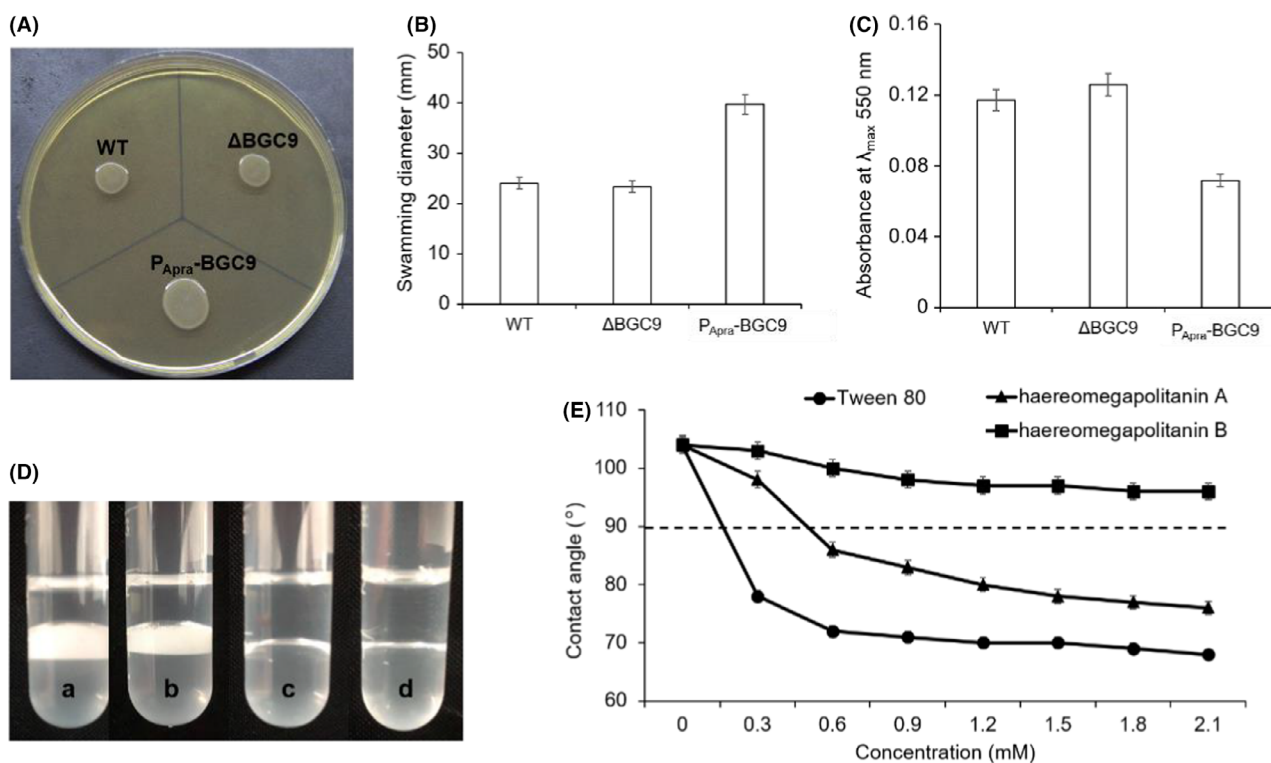


Fig. 4. Biological functions of lipopeptide haereomegapolitanins.

A. Swarming phenotypes of wild type (WT) as well as the BGC9 activated mutant (P_{Apra}-BGC9) and BGC9 inactivated mutant (ΔBGC9) of *P. megapolitana* DSM 23488.

B. Swarming diameter.

C. Biofilm formation.

D. Emulsification activity of culture supernatants. (a) Tween-80; (b) haereomegapolitanin A; (c) haereomegapolitanin B. (d) H₂O.

E. Contact angle measurement of aqueous solutions of lipopeptides. Tween 80 is a positive control.

mediators is known. Previous research has revealed that *P. megapolitana* isolated from mosses show antifungal activity (Eberl and Vandamme, 2016), and the antifungal compounds of *P. megapolitana* have not been identified. Our results indicate that haereomegapolitanin A promotes the swarming of bacteria and has biosurfactant activity; thus, haereomegapolitanin A probably promotes the colonization of bacteria. Our researches may bring new directions to the study of bacterial colonization infection.

Due to the low yields and various derivatives of the BGC5 products, we failed to obtain enough pure products for NMR recording to determine their structures. We tested the surfactant activity of the crude extract of BGC5 activated mutant (P_{Apra}-BGC5). The results showed the products of BGC5 did not have surfactant property (Fig. S13). The function and structures also need to be investigated in the future.

In summary, we have demonstrated that a pair of recombinases from Burkholderiales strain DSM 7029 is effective in the genome modification of *P. megapolitana* DSM 23488. New metabolites were also identified through promoter knock-in aided by heterologous

recombineering system. This recombineering-mediated genome editing system offers more possibility for manipulation of more Burkholderiales species, facilitating secondary metabolite mining and accelerating the discovery of potentially novel bioactive compounds.

Acknowledgements

This work was supported by the National Key R&D Programme of China (2017YFD0201400, 2019YFA0905700), National Natural Science Foundation of China (31670098, 31670097), the Shandong Provincial Natural Science Foundation, China (ZR2019JQ11, ZR2017BC059), the China Postdoctoral Science Foundation (2019M652373), the Recruitment Programme of Global Experts (Y.Z.); the Qilu Youth Scholar Startup Funding of SDU (to X.B.). We would like to thank Drs. Zhifeng Li and Jingyao Qu from State Key laboratory of Microbial Technology of Shandong University for help and guidance in LC-MS.

Conflict of interest

None declared.

References

- Bachmann, B.O., Van Lanen, S.G., and Baltz, R.H. (2014) Microbial genome mining for accelerated natural products discovery: is a renaissance in the making? *J Ind Microbiol Biotechnol* **41**: 175–184.
- Biggins, J.B., Gleber, C.D., and Brady, S.F. (2011) Acyldepsipeptide HDAC inhibitor production induced in *Burkholderia thailandensis*. *Org Lett* **13**: 1536–1539.
- Blin, K., Medema, M.H., Kottmann, R., Lee, S.Y., and Weber, T. (2017) The antiSMASH database, a comprehensive database of microbial secondary metabolite biosynthetic gene clusters. *Nucleic Acids Res* **45**: D555–D559.
- Cimercancic, P., Medema, M.H., Claesen, J., Kurita, K., Wieland Brown, L.C., Mavrommatis, K., *et al.* (2014) Insights into secondary metabolism from a global analysis of prokaryotic biosynthetic gene clusters. *Cell* **158**: 412–421.
- Depoorter, E., Bull, M.J., Peeters, C., Coenye, T., Vandamme, P., and Mahenthiralingam, E. (2016) *Burkholderia*: an update on taxonomy and biotechnological potential as antibiotic producers. *Appl Microbiol Biotechnol* **100**: 5215–5229.
- Eberl, L., and Vandamme, P. (2016) Members of the genus *Burkholderia*: good and bad guys. *F1000Research* **5**: 1007.
- el-Banna, N. and Winkelmann, G. (1998) Pyrrolnitrin from *Burkholderia cepacia*: antibiotic activity against fungi and novel activities against streptomycetes. *J Appl Microbiol* **85**: 69–78.
- Esmaeel, Q., Pupin, M., Kieu, N.P., Chataigne, G., Bechet, M., Davel, J., *et al.* (2016) *Burkholderia* genome mining for nonribosomal peptide synthetases reveals a great potential for novel siderophores and lipopeptides synthesis. *Microbiologyopen* **5**: 512–526.
- Esmaeel, Q., Pupin, M., Jacques, P., and Leclere, V. (2018) Nonribosomal peptides and polyketides of *Burkholderia*: new compounds potentially implicated in biocontrol and pharmaceuticals. *Environ Sci Pollut Res Int* **25**: 29794–29807.
- Fu, J., Bian, X., Hu, S., Wang, H., Huang, F., Seibert, P.M., *et al.* (2012) Full-length RecE enhances linear-linear homologous recombination and facilitates direct cloning for bioprospecting. *Nat Biotechnol* **30**: 440–446.
- Hertweck, C. (2009) The biosynthetic logic of polyketide diversity. *Angew Chem Int Ed Engl* **48**: 4688–4716.
- Huo, L., Hug, J.J., Fu, C., Bian, X., Zhang, Y., and Müller, R. (2019) Heterologous expression of bacterial natural product biosynthetic pathways. *Nat Prod Rep* **36**: 1412–1436.
- Kang, Y., Norris, M.H., Wilcox, B.A., Tuanyok, A., Keim, P.S., and Hoang, T.T. (2011) Knockout and pullout recombineering for naturally transformable *Burkholderia thailandensis* and *Burkholderia pseudomallei*. *Nat Protoc* **6**: 1085–1104.
- Liu, X., and Cheng, Y.Q. (2014) Genome-guided discovery of diverse natural products from *Burkholderia* sp. *J Ind Microbiol Biotechnol* **41**: 275–284.
- Liu, X., Xie, F., Doughty, L.B., Wang, Q., Zhang, L., Liu, X., and Cheng, Y.Q. (2018) Genomics-guided discovery of a new and significantly better source of anticancer natural drug FK228. *Synth Syst Biotechnol* **3**: 268–274.
- Medema, M.H., Kottmann, R., Yilmaz, P., Cummings, M., Biggins, J.B., Blin, K., *et al.* (2015) Minimum information about a biosynthetic gene cluster. *Nat Chem Biol* **11**: 625–631.
- Mukherjee, S., Seshadri, R., Varghese, N.J., Eloie-Fadrosch, E.A., Meier-Kolthoff, J.P., Goker, M., *et al.* (2017) 1,003 reference genomes of bacterial and archaeal isolates expand coverage of the tree of life. *Nat Biotechnol* **35**: 676–683.
- Nakayama, M., and Ohara, O. (2005) Improvement of recombination efficiency by mutation of red proteins. *Biotechniques* **38**: 917–924.
- Raaijmakers, J.M., De Bruijn, I., Nybroe, O., and Ongena, M. (2010) Natural functions of lipopeptides from *Bacillus* and *Pseudomonas*: more than surfactants and antibiotics. *FEMS Microbiol Rev* **34**: 1037–1062. <https://doi.org/10.1111/j.1574-6976.2010.00221.x>.
- Rutledge, P.J., and Challis, G.L. (2015) Discovery of microbial natural products by activation of silent biosynthetic gene clusters. *Nat Rev Microbiol* **13**: 509–523.
- Sussmuth, R.D., and Mainz, A. (2017) Nonribosomal peptide synthesis-principles and prospects. *Angew Chem Int Ed Engl* **56**: 3770–3821.
- Tang, B., Yu, Y., Liang, J., Zhang, Y., Bian, X., Zhi, X., and Ding, X. (2019) Reclassification of 'Polyangium brachysporum' DSM 7029 as *Schlegelella brevitalia* sp. nov. *Int J Syst Evol Microbiol* **69**: 2877–2883.
- Thongdee, M., Gallagher, L.A., Schell, M., Dharakul, T., Songsivilai, S., and Manoil, C. (2008) Targeted mutagenesis of *Burkholderia thailandensis* and *Burkholderia pseudomallei* through natural transformation of PCR fragments. *Appl Environ Microbiol* **74**: 2985–2989.
- Thongkongkaew, T., Ding, W., Bratovanov, E., Oueis, E., Garci, A.A.M.A., Zaburannyi, N., *et al.* (2018) Two types of threonine-tagged lipopeptides synergize in host colonization by pathogenic *Burkholderia* species. *ACS Chem Biol* **13**: 1370–1379.
- Vandamme, P., Opelt, K., Knochel, N., Berg, C., Schonmann, S., De Brandt, E., *et al.* (2007) *Burkholderia bryophila* sp. nov. and *Burkholderia megapolitana* sp. nov., moss-associated species with antifungal and plant-growth-promoting properties. *Int J Syst Evol Microbiol* **57**: 2228–2235.
- Wang, C., Henkes, L.M., Doughty, L.B., He, M., Wang, D., Meyer-Almes, F.J., and Cheng, Y.Q. (2011) Thailandepsins: bacterial products with potent histone deacetylase inhibitory activities and broad-spectrum antiproliferative activities. *J Nat Prod* **74**: 2031–2038.
- Wang, X., Zhou, H., Chen, H., Jing, X., Zheng, W., Li, R., *et al.* (2018) Discovery of recombinases enables genome mining of cryptic biosynthetic gene clusters in *Burkholderiales* species. *Proc Natl Acad Sci USA* **115**: E4255–E4263.
- Yin, J., Zhu, H., Xia, L., Ding, X., Hoffmann, T., Hoffmann, M., *et al.* (2015) A new recombineering system for *Photorhabdus* and *Xenorhabdus*. *Nucleic Acids Res* **43**: e36.
- Yin, J., Zheng, W., Gao, Y., Jiang, C., Shi, H., Diao, X., *et al.* (2019) Single-stranded DNA-binding protein and exogenous RecBCD inhibitors enhance phage-derived homologous recombination in *Pseudomonas*. *iScience* **14**: 1–14.

Zarins-Tutt, J. S., Barberi, T. T., Gao, H., Mearns-Spragg, A., Zhang, L., Newman, D. J., and Goss, R. J. (2016) Prospecting for new bacterial metabolites: a glossary of approaches for inducing, activating and upregulating the biosynthesis of bacterial cryptic or silent natural products. *Nat Prod Rep* **33**: 54–72.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Fig. S1. Diagram for construction, verification and metabolic analysis of BGC5 activated mutant and inactivated mutant in DSM 23488.

Fig. S2. Diagram for construction, verification and metabolic analysis of BGC9 activated mutant and inactivated mutant in DSM 23488.

Fig. S3. Diagram for construction, verification and metabolic analysis of *orf1-hrmB* deletion in DSM 23488- P_{Apra} -BGC9.

Fig. S4. ^1H NMR spectrum (500 MHz) of haereomegapolitanin A (**1**) in $\text{DMSO-}d_6$.

Fig. S5. ^{13}C NMR spectrum (125 MHz) of haereomegapolitanin A (**1**) in $\text{DMSO-}d_6$.

Fig. S6. DEPT spectrum of haereomegapolitanin A (**1**) in $\text{DMSO-}d_6$.

Fig. S7. HSQC spectrum of haereomegapolitanin A (**1**) in $\text{DMSO-}d_6$.

Fig. S8. $^1\text{H-}^1\text{H}$ COSY spectrum of haereomegapolitanin A (**1**) in $\text{DMSO-}d_6$.

Fig. S9. HMBC spectrum of haereomegapolitanin A (**1**) in $\text{DMSO-}d_6$.

Fig. S10. Comparative analysis of MS/MS fragmentation of haereomegapolitanin A (**1**) and haereomegapolitanin B (**2**).

Fig. S11. Marfey's analysis of the amino acid constituents of haereomegapolitanin A (**1**) by LC-MS.

Fig. S12. Complete structure with carbon number (**A**), and key COSY and HMBC correlations of **1** (**B**).

Fig. S13. Biological functions for BGC5 activated products.

Table S1. Predicted BGCs in *P. megapolitana* DSM 23488 by antiSMASH 5.0.

Table S2. Predicted gene function and transcript level (fpkm) of BGC5 and BGC9 of DSM 23488.

Table S3. Strains, plasmids and mutants in this work.

Table S4. Primers used in this study.

Table S5. The ^1H (500 MHz) and ^{13}C NMR (125 MHz) data of haereomegapolitanin A (**1**) in $\text{DMSO-}d_6$.