An integrated workflow for phenazine-modifying enzyme characterization

R. Cameron Coates[1] · Benjamin P. Bowen[1,2] · Ernst Oberortner[1] · Linda Thomashow[3,4] · Michalis Hadjithomas[1] · Zhiying Zhao[1] · Jing Ke[1] · Leslie Silva[1] · Katherine Louie[1] · Gaoyan Wang[1] · David Robinson[1] · Angela Tarver[1] · Matthew Hamilton[1] · Andrea Lubbe[2] · Meghan Feltcher[5] · Jeffery L. Dangl[5,6,7,8,9] · Amrita Pati[1] · David Weller[3,4] · Trent R. Northen[1,2] · Jan-Fang Cheng[1,2] · Nigel J. Mouncey[1] · Samuel Deutsch[1,2,10] · Yasuo Yoshikuni[1,2,10]

Abstract

Increasing availability of new genomes and putative biosynthetic gene clusters (BGCs) has extended the opportunity to access novel chemical diversity for agriculture, medicine, environmental and industrial purposes. However, functional characterization of BGCs through heterologous expression is limited because expression may require complex regulatory mechanisms, specific folding or activation. We developed an integrated workflow for BGC characterization that integrates pathway identification, modular design, DNA synthesis, assembly and characterization. This workflow was applied to characterize multiple phenazine-modifying enzymes. Phenazine pathways are useful for this workflow because all phenazines are derived from a core scaffold for modification by diverse modifying enzymes (PhzM, PhzS, PhzH, and PhzO) that produce characterized compounds. We expressed refactored synthetic modules of previously uncharacterized phenazine BGCs heterologously in *Escherichia coli* and were able to identify metabolic intermediates they produced, including a previously unidentified metabolite. These results demonstrate how this approach can accelerate functional characterization of BGCs.

Keywords Synthetic biology · Biosynthesis · Phenazine · Pathway · Refactored · Pathway design

✉ Yasuo Yoshikuni
yyoshikuni@lbl.gov

[1] US DOE Joint Genome Institute, Walnut Creek, CA, USA

[2] Environmental Genomics and Systems Biology Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA

[3] USDA Agricultural Research Service, Wheat Health, Genetics and Quality, Washington State University, Pullman, WA, USA

[4] Department of Plant Pathology, Washington State University, Pullman, WA, USA

[5] Department of Biology, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

[6] Howard Hughes Medical Institute, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

[7] Curriculum in Genetics and Molecular Biology, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

[8] Department of Microbiology and Immunology, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

[9] Carolina Center for Genome Sciences, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

[10] Biological Systems and Engineering Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA

Introduction

Genome sequencing efforts have greatly expanded the identification of biosynthetic gene clusters (BGCs) from diverse bacteria, fungi and plants. These BGCs have enormous potential to produce novel metabolites which may be used for agricultural, medicinal, environmental or industrial applications [20]. However, most of these BGCs exhibit complex regulation of pathway expression within their native hosts. Although there are many tools and methods available for cloning BGCs along with well-characterized inducible promoters, these approaches are often not sufficient to remove complex native pathway regulation and activate functional expression of BGCs for characterization of their products. Additionally, codon- and GC-bias can limit reliable expression in heterologous bacterial hosts. More recently, a synthetic biology approach leveraging codon optimization, systematic refactoring and DNA synthesis was proposed as an effective strategy to more reliably express BGCs, opening a new paradigm in novel metabolite research [18, 20].

We chose to study phenazine BGCs as model compounds for this investigation. Phenazines (Phz) are a family of compounds sharing a dibenzo-annulated pyrazine as a core structure. Pierson et al. [15] first discovered a core phenazine BGC comprising *phz*A-G genes in *Pseudomonas aureofaciens,* and subsequently many other *Pseudomonads.* These pathways have been extensively characterized [1, 9, 16]. The *phz*A-G genes are commonly found in phenazine-producing bacteria [11] and are responsible for producing phenazine-1-carboxylic acid (PCA) [19, 23] (Fig. 2). Recent bioinformatics analyses of diverse phenazine BGCs revealed the modular nature of these clusters [12, 19] and suggest that greater diversity among associated modules can provide phenazines with various chemical, physical, and biological properties. This architecture also facilitates stepwise optimization and functional characterization of these BGCs, making them ideal models for synthetic biology investigation.

Almost all phenazines are known to exhibit broad antibacterial and antifungal bioactivity due to their redox activity [11]. The phenazines produced in the rhizosphere inhibit soil-borne fungal pathogens and contribute to the suppression of "take-all" disease of wheat in the Pacific Northwest, United States and *Fusarium* wilt in the Châteaurenard region of France [23]. Phenazine production among biocontrol strains contributes to multiple facets of the process of disease suppression including rhizosphere competence and biofilm formation by the bacteria [16] and ultimately, to pathogen inhibition [23]. Generally, the activity and specificity of phenazines in antifungal interactions is related to their redox potential, solubility, and stability in each environment. For example, methoxy phenazines are more susceptible to oxidation, while PCA is less susceptible [22]. Soil pH can also significantly impact the bioactivity of phenazines. PCA rapidly loses its bioactivity above pH 5, while phenazine-1-carboxamide (PCN) is not influenced by soil pH and, therefore, maintains bioactivity over a broad range of soil conditions [3]. Therefore, this synthetic biology approach may allow for production of designed phenazines with more suitable properties for agricultural and medicinal uses.

We used DOE JGI's IMG/ABC (Department of Energy Joint Genome Institute, Integrated Microbial Genomes/ Atlas of Biosynthetic Clusters) along with the *phz*A-G operon and its associated modules to identify over 1000 phenazine BGCs. Among them, 26 phenazine BGCs with unique pathway architectures were identified [7]. The *phz*A-G operon and homologs of known phenazine-modifying enzymes including *phzS*, *phzH*, and *phzM* were identified and labeled accordingly (Figs. 1, 2). Interestingly, many of these modules reside in genera not previously known to produce phenazines. A few BGCs containing unique modular enzymes were further selected, with preference given to sources with a likely role in microbe– and plant–microbe interactions. Here, we expressed representative accessory
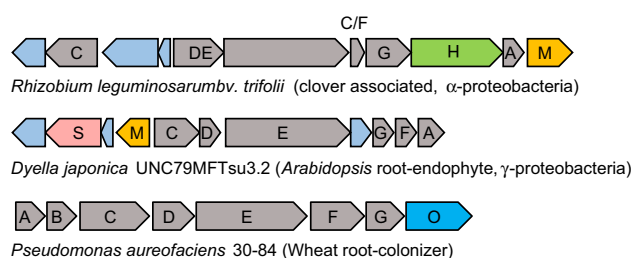
Fig. 1 The phenazine BCGs selected for characterization. Genetic architecture of three BGCs selected for this study is shown. Approximately, 1000 phenazine BGCs with diverse accessory modules were found through the analyses of over 25,000 isolate genome sequences available through the JGI IMG. Further clustering analyses of these phenazine BGCs revealed 26 BGCs with unique genetic architectures [7]. Finally, these BGCs were selected because they were found in plant-associated microbes, which may suggest their involvement in microbe–plant interactions

modular enzymes of these phenazine pathways along with the core phenazine gene cluster using a modular design and assembly strategy. This approach enabled us to explore a wide diversity of phenazine structures and pathways.

Materials and methods

Microbes and reagents

All cloning work was performed using *Escherichia coli* DH10B, DH5a, or Transformax™ EC100D™ pir+ as the host strains. *E. coli* BL21(DE3) was used as a production host for phenazines. The phenazine BGCs are identified from the strains listed in Table 1. Each strain was cultured and used for phenazine production investigations. All reagents were purchased from either Thermo Fisher Scientific or MilliporeSigma, unless otherwise stated. Phenazine-1-carboxylic acid (PCA) and 1-hydroxyphenazine (1-OHPhz) were purchased from Tractus Chemical. Phenazine-1-carboxamide (PCN) and pyocyanin (PYO) were purchased from Santa Cruz Biotechnology.

Construction of a core Phz-BGC for PCA production

A phenazine operon comprising *phzA-G* from *Pseudomonas synxantha 2-79* was previously characterized to produce PCA by Mavrodi et al. [9]. This core PCA pathway was cloned from a plasmid provided by Dr. Thomashow's group. The operon was cloned under the control of a T7 Phi10 promoter via LacI into a pBeloBAC vector using Gibson assembly [26]. The resulting construct (pW7_PCA) was subsequently sequence verified. The primers used for PCR are listed on Table 2.
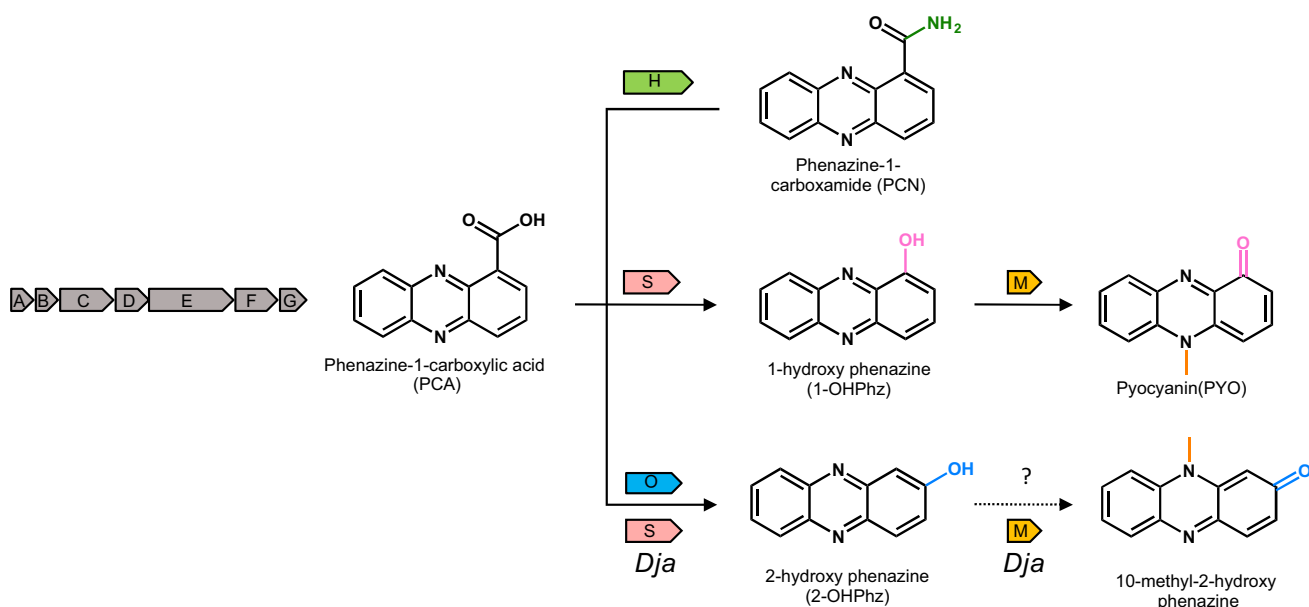
Fig. 2 Chemical reactions catalyzed by known PCA-modifying enzymes. PCA is produced by the core phenazine BGC. PhzH catalyzes transamidation to form PCN. PhzS catalyzes reductive decarboxylation and hydroxylation at the 1-position to form 1-OHPhz. PhzM then further modifies 1-OHPhz to form PYO. PhzO was identified in *P. aureofaciens* 30-84 and catalyzes reductive hydroxylation to form 2-OHPCA followed by spontaneous decarboxylation to form 2-OHPhz. In this study, a synthetic module from *D. japonica* produced a PYO isomer. Because there are only four possible structural variations for PYO, and two isomers of PYO, PYO and 10-methyl-2-OHPhz, are likely more stable than the other isomers. We speculate that the product is 10-methyl-2-OHPhz (NMR verification is required)

Table 1 Strains investigated in this study with isolation source and class

Strain	Isolation source	Class
Pseudomonas synxantha 2-79	Wheat root-colonizer	ϒ-Proteobacteria
Pseudomonas aureofaciens 30-84	Wheat root-colonizer	ϒ-Proteobacteria
Dyella japonica UNC79MFTsu3.2	*Arabidopsis* root-endophyte	ϒ-Proteobacteria
Rhizobium leguminosarum bv. *trifolii* CC283b	Clover associated	α-Proteobacteria

**Table 2** Primers used in this experiment

Primer	Sequence	Plasmid assembled
SBP0463_PCA_pApr_1F	aaataattttgtttaactttaagaaggagaacactgcctatgcccgg	pW5_PCA
SBP0624_pKm_PhzO_Phi9F	gttttccgggacgccggctgacggcacggccacgcgtttaaaccgcc	pW7_PhzO
SBP0613_PhzSM_Djapon_F	tctaaatccaagacattcacccaacaggcatatc	pW7_PhzSM_Djapon
SBP0614_PhzHM_RhizoF	tcttacgactttccattattcaagagggttag	pW7_PhzHM_Rhizo
SBP0464_PCA_pApr_R	accgcctttgagtgagcaggtccgccggcgttggacgagcgtcagggct-gcaaacgccgc	pW5_PCA, pW7_PhzO, pW7_PhzSM_Djapon, pW7_PhzHM_Rhizo, pW7_PhzH_Balni, pW7_PhzS_Patro
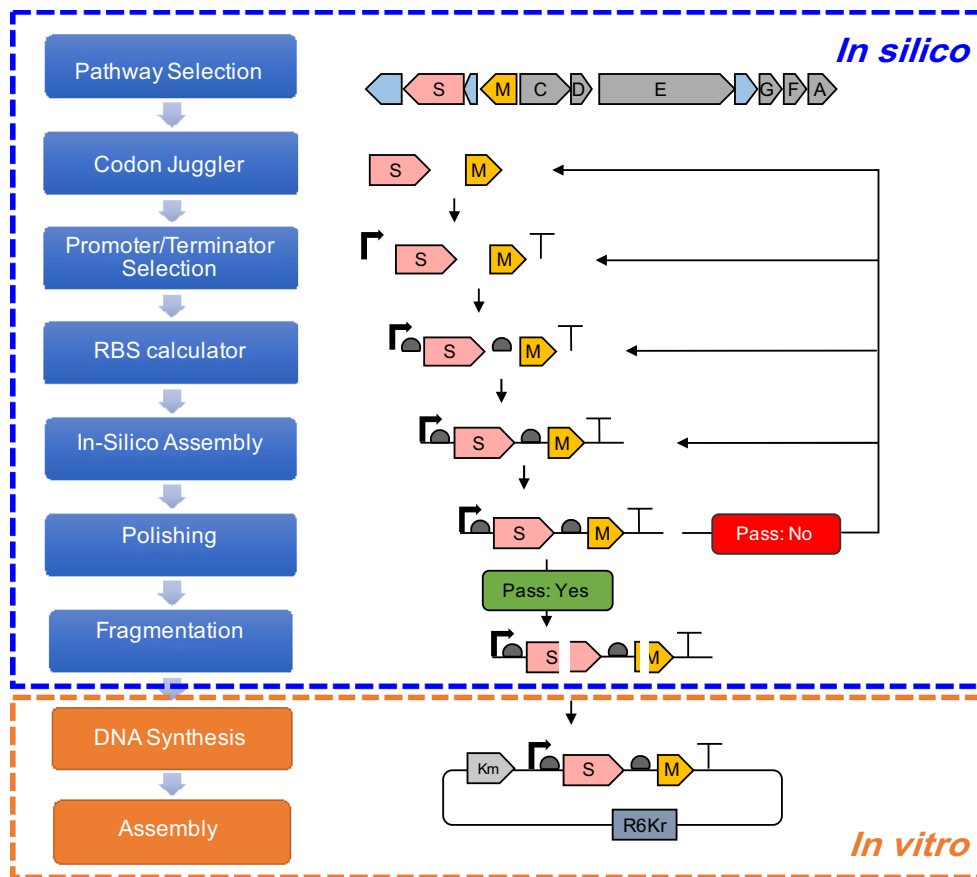
Fig. 3 An integrated workflow for BGC characterization. The integrated workflow for BGC characterization is shown. The workflow comprises two phases, in silico design and assembly and in vitro assembly. In the in silico phase, genes of interest from the selected BGCs are extracted and are codon optimized (Codon juggler). A promoter and a terminator are selected (promoter/terminator selection), and ribosome-binding site (RBS) with target expression level is cal-culated (RBS calculation). These parts are assembled into a construct in silico (in silico assembly), all sequences that are not favorable for de novo DNA synthesis are replaced (polishing), and the construct is fragmented into shorter fragments to facilitate the synthesis (fragmentation). These DNA fragments are subsequently synthesized and assembled into a construct (in vitro DNA synthesis and assembly)

Fig. 4 Four synthetic phenazine BGC constructs tested. The architecture of six synthetic phenazine BGCs is shown. The *phzH, M, S,* and *O* gene homologs from four phenazine BGCs described in Fig. 1 are refactored, synthesized, and assembled with the PCA-producing core phenazine BGC from *P. synxantha 2-79*
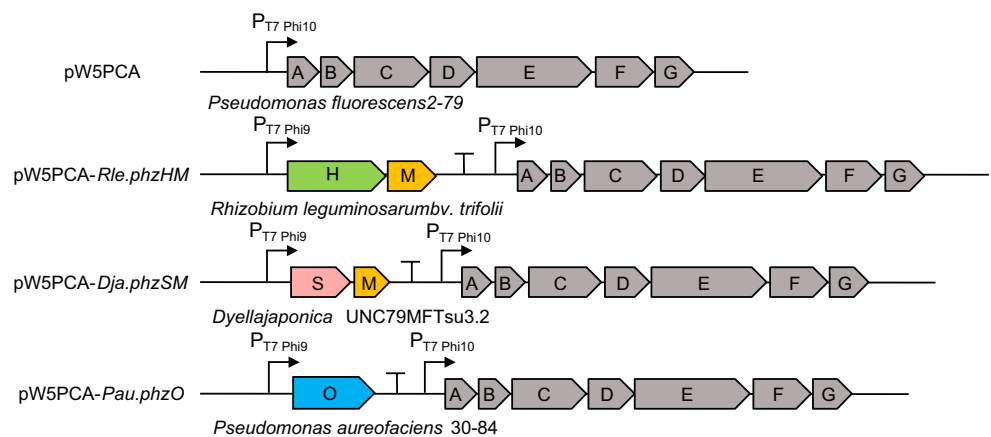
BGC refactoring, synthesis, and assembly

The modules of PCA-modifying enzymes were designed for synthesis and assembly using an integrated workflow as illustrated in Fig. 3, and the design for each BGC is illustrated in Fig. 4. The process involves in silico design and in vitro assembly.

In silico design

In the "Codon Juggler" step, we used Build-Optimization Software Tools (BOOST) [13] to optimize codon usage of the coding sequences for our target host strain *E. coli*. The codon usage table that was used for codon optimization was derived from *Pseudomonas putida* KT2440. A balanced codon usage was selected for codon optimization. In the "Promoter/Terminator Selection" step, we selected the promoters for the core phenazine BGC and the phenazine-modifying modules. The core phenazine BGC was designed to be expressed using the T7 Phi10 promoter. A T7 promoter variant Phi 9 was selected for expression of phenazine-modifying enzymes from the T7 bacteriophage genome (accession: NC_001604) by searching for the 19-bp T7 consensus sequence (TAATACGACTCACTATAG G) and was previously described in the literature as being among strong T7 promoters designed for inducible expression via IPTG in strains harboring T7 RNA polymerase such as *E. coli* BL21(DE3) [24, 25]. While a terminator was not used for *phzA-G*, a T7 terminator was used for expression of phenazine-modifying enzymes. In the "Ribosome Binding Site (RBS) Calculator" step, we calculated the sequence of the RBS for each codon-optimized coding sequence. We utilized the RBS calculator v2.0 [6] to design the synthetic RBS sequences to achieve a targeted translation initiation rate of 10,000 in the context of each coding sequence and the Phi 9 T7 promoter selected for expression.

In the "In silico assembly" step, we used j5 integrated into a design, implementation and verification automation (DIVA) platform to perform in silico assembly of the parts into the designed BGCs and their operons [2, 8]. In the "Polishing" step, sequences that were difficult to undergo de novo synthesis (global/local high/low GC contents, homopolymers, and repeats) were eliminated and modified accordingly. If difficult sequences were identified in coding regions, the codon juggler step was repeated and reanalyzed. If difficult sequences were identified in RBS sites, the RBS calculator was rerun and reanalyzed. If difficult sequences were identified in promoter or terminators, alternative promoters and terminators were selected and reanalyzed. Finally, in the "Fragmentation" step, each construct was partitioned into shorter fragments (in general, 1.5–3 kb) to facilitate de novo synthesis and assembly. We utilized BOOST to perform the "Polishing" and the "Fragmentation" steps in an automated manner.

While this workflow can generate numbers of design variants, we decided to focus on a single design for each pathway. DNA fragments (length 1.5–3 Kb) were then ordered as dsDNA parts from SGI-DNA and subsequently PCR amplified and assembled via gibson assembly [26] into accessory vectors (described below) to be used for recombineering pathway assembly.

In vitro assembly

Refactored modules were assembled into accessory plasmids (R6K ori) that were used to assemble final expression plasmids via a recombineering technology [4, 21] (Fig. 4). Phenazine expression plasmids were maintained in DH10B host strains that also contained the λ-red recombination machinery on pKD46 [4]. The λ-red recombinase expression was induced with 2% arabinose at OD600 of 0.1 and cultured for 4 h. Competent cells were prepared following standard procedures [4], and PCR amplicons containing 250 bp upstream and 40 bp downstream homology were transformed into DH10B host strains harboring vectors containing the core Phz operon and pKD46. Counter-selection on kanamycin and apramycin was performed on resulting transformants. Transformants that exhibited a counter-selection pattern consistent with successful insertion of PCA-modifying modules and their adjacent marker (apramycin) were subsequently sequence verified.

Microbial production of Phz metabolites

Escherichia coli BL21(DE3) harboring plasmids with the synthetic *phz* operon (Fig. 4) were tested for Phz production, inoculated into LB from glycerol stocks and grown overnight to prepare seed cultures, and used to inoculate fresh LB at an OD600 of 0.1. To induce expression of the phenazine BGCs, isopropyl-1-thio-β-D-galactopyranoside (IPTG) was added to each culture at 1.0 mM as a final concentration when cultures reached an OD600 of between 0.5 and 0.7. These cultures were further grown at 37 °C for 4 h and harvested via centrifugation at 3200 rpm for 5 min. A portion of the supernatant (200 μl) was removed and extracted by mixing with 200 μl of 1.25% trifluoroacetate and then adding 800 μl ethyl acetate. Samples were then centrifuged at 3200 rpm for 10 min. The top 400 μl of the top phases containing ethyl acetate was then removed from each extract and filtered with a 0.2-μm filter. Extracts were then dried in a speed vacuum for 1 h and suspended in 50 μl of acetonitrile.

LCMS analysis

High-resolution electrospray ionization mass spectrometry analyses were performed on each sample extract using an Agilent 1290 LC stack for ultrahigh performance liquid chromatography (UHPLC) coupled to a Q-Exactive (Thermo Fisher Scientific, San Jose, CA, USA) to collect MS1 and MS2 data (we refer to the system as LC–MS hereafter) and a photodiode array detector (Agilent Technology, Santa Clara, CA, USA) to collect UV–Vis spectra. Each sample extract resuspended in 1 µl of acetonitrile was injected onto a C18 column (Agilent ZORBAX Eclipse Plus C18, Rapid Resolution HD, 2.1 × 50 mm, 1.8 µm) held at 60 °C. Chromatography was performed at a flow rate of 0.4 ml/min, with the C18 column equilibrated with 100% buffer A (100% LC–MS water w/0.1% formic acid) for 1 min, diluting buffer A down to 0% with buffer B (100% ACN w/0.1% formic acid) over 8 min, and followed by 1.5 min of isocratic elution with 100% B. Full MS spectra were collected in positive ion mode from 80 to 1200 *m/z* at 70,000 resolution, and MS2 fragmentation data acquired using 10, 20 and 30 eV collision energies (averaged) at 17,500 resolution. Sample data were analyzed using Metatlas (https://metatlas.nersc.gov/), Xcalibur (Thermo Fisher Scientific, San Jose, CA, USA), and Mzmine (http://mzmine.github.io/) software. Standards for PCA, PCN, PYO, and 1-OHPhz were used for characterization of compounds produced from microbial systems.

Results

Selection of candidate phenazine BGCs for characterization

Hadjithomas et al. recently reported development of the computational platform, an atlas of biosynthetic gene clusters (ABC) within the integrated microbial genomes (IMG) system (https://img.jgi.doe.gov/abc), which enables systematic discovery of novel metabolite BGCs [7]. The IMG/ABC contains two classes of objects describing BGCs and known secondary metabolites associated with those BGCs. Using this platform, Hadjithomas et al. analyzed more than 25,000 isolate genomes available through IMG and identified over 1000 phenazine BCGs with diverse accessory modules [7]. These phenazine clusters were further analyzed and classified based on their architectural similarity, which resulted in the identification of 26 representative phenazine modules with unique architectures [7]. Interestingly, many of the identified phenazine BGCs were found in genera not previously known to produce phenazines. Further, we selected three phenazine BCGs based on the likelihood of their involvement in plant–microbe association (Fig. 1).

Next, we investigated whether each BGC contains orthologues to any genes of which function was previously characterized, including *phzS*, *phzO*, *phzH* and *phzM* (Fig. 2, Table 3). These genes encode enzymes capable of modifying PCA. The *phzS* and *phzM* genes were originally found in *Pseudomonas aeruginosa* PAO1. PhzS catalyzes reductive decarboxylation and hydroxylation of 1-position to yield 1-OHPhz. PhzM catalyzes methylation of the N5-position of 1-OHPhz to form PYO [10], which is involved in anaerobic energy generation of *P. aeruginosa* while it is forming biofilms [17]. *P. aeruginosa* PAO1 is also known to produce PCN [11], catalyzed by PhzH. The *phzH* gene is also found in *P. chlororaphis* [3]. To identify target genes for functional characterization, we identified homologs of known phenazine-modifying genes. Homology of the target genes selected for characterization is listed in Table 3 with percent identity and similarity to known phenazine-modifying genes (Fig. 1, Table 3).

Additionally, we included the gene *phzO* from *P. aureofaciens* 30–84 as it catalyzes a different reaction from the above-mentioned three enzymes. PhzO catalyzes reductive hydroxylation at 2-position to yield 2-OH-PCA which is spontaneously decarboxylated to yield 2-OHPhz [5].

Development of strains for PCA production

To characterize the function of *phzS, phzM,* and *phzH* gene homologs in the selected BGCs, we first built strains capable of supplying PCA as a substrate. In vivo production of PCA in *E. coli* has been previously demonstrated [19]. For the development of the Phz-producing strain, we cloned the complete *phz* core BGC (*phzA-G*) from *P. synxantha 2-79* [1, 9]. We confirmed PCA production from the resulting strain (Fig. 5a–c).

Table 3 Percent identity and similarity to the PCA-modifying enzymes

Gene name	Sources (strain)	Orthologue	Identity (%)	Similarity (%)
*phz*S	*D. japonica UNC79MFTsu3.2*	*phz*S *P. aeruginosa* PAO1	47.4	63.3
*phz*S	*D. japonica UNC79MFTsu3.2*	*phz*O *P. chlororaphis* subsp. *aureofaciens*	13.9	29.4
*phz*M	*D. japonica UNC79MFTsu3.2*	*phz*M *P. aeruginosa* PAO1	30.2	48.3
*phz*H	*Rhizobium leguminosarum bv. trifolii CC283b*	*phz*H *P. chlororaphis* PCL1391	44.9	62.8
*phz*M	*Rhizobium leguminosarum bv. trifolii CC283b*	*phz*M *P. aeruginosa* PAO1	42.4	57.3
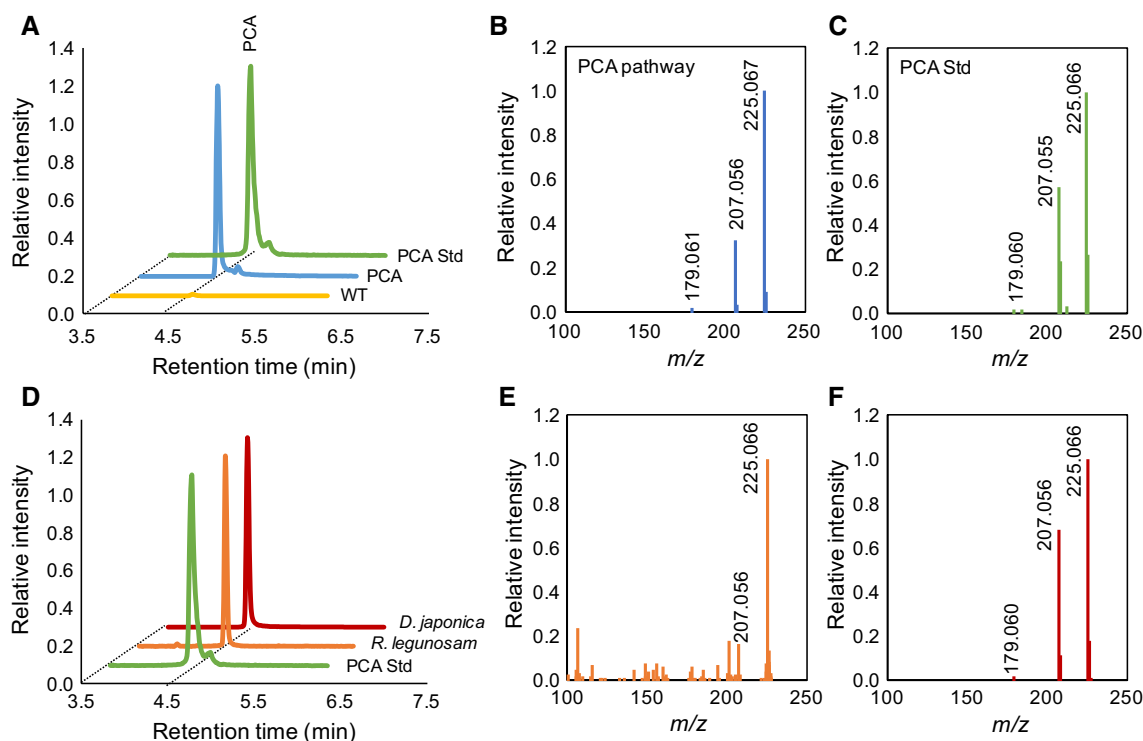
Fig. 5 Production of PCA. PCA production from an *E. coli* BL21(DE3) strain harboring the core PCA-producing BGC are shown: the LC–MS chromatograms of *m/z* 225.06 (**a**) and the MS2 spectra of peaks with *m/z* 225.06 [M+H]$^+$ of the culture extract of *E. coli* with the core phenazine BGC (**b**) and PCA chemical standard (**c**). PCA production from *R. leguminosarum* and *D. japonica* are shown: the LC–MS chromatograms of *m/z* 225.06 (**d**) and the MS2 spectra of *m/z* 225.06 [M+H]$^+$ from the culture extracts of *R. leguminosarum* (**e**) and *D. japonica* (**f**). In these panels, the culture extracts of *E. coli* BL21(DE3) harboring pBeloBAC and pW5PCA are shown

in light orange and blue, respectively. The PCA chemical standard is shown in green. The culture extracts of *R. leguminosarum* and *D. japonica* are shown in dark orange and red. The MS2 spectrum of PCA for the culture extract of *R. leguminosarum* was very noisy, because the PCA production level was nearly 1000-fold lower than those of other strains. However, we still detected signature fragment ions such as *m/z* 207.056 and 197.061 on the spectrum. MS2 spectra collected at 30 eV collision energy. Each LC–MS trace/MS2 spectrum is a representative of triplicate experiments (color figure online)

PCA production from native strains

To investigate whether the native strains carrying the bioinformatically selected Phz BGCs can produce any phenazine derivative molecules, we obtained *Rhizobium leguminosarum* from Dr. Wayne Reeve's group at Murdoch University and *Dyella japonica* from Jeffrey Dangl's group at the University of North Carolina Chapel Hill. We grew *R. leguminosarum* in tryptone yeast extract medium and *D. japonica* was grown in LB medium. Both strains were grown at 28 °C in 50-ml flasks for 72 h. The extracts of supernatants prepared as described in materials and methods section were analyzed by the LC–MS. We confirmed PCA production from these strains through comparison with chemical standard (Fig. 5d–f).

## Characterization of *phzS, phzM, and phzH* homologs in the selected BGCs

We refactored, synthesized and assembled modular BGCs comprising putative PCA-modifying enzymes that are similar to *phzS, phzM,* and *phzH* genes in each BGC (Fig. 4). These modules were assembled along with the PCA-producing core phenazine BGC. All four synthetic BGCs were transformed and expressed in *E. coli* BL21(DE3). The extracts of supernatants prepared as described in materials and methods section were analyzed by LC–MS. We verified production of Phz derivatives in strains expressing these synthetic BGCs in comparison with chemical standards (Figs. 6, 7).

The module derived from *R. leguminosarum* comprises *phzH* and *phzM* homologs. While we expected to see 5- or 10-methyl PCN, we only detected a peak corresponding to PCN (Figs. 2, 6). The PCN product matched the LC retention time, exact mass, and MS2 spectra of the high-resolution PCN chemical standard (Fig. 6). The expected product
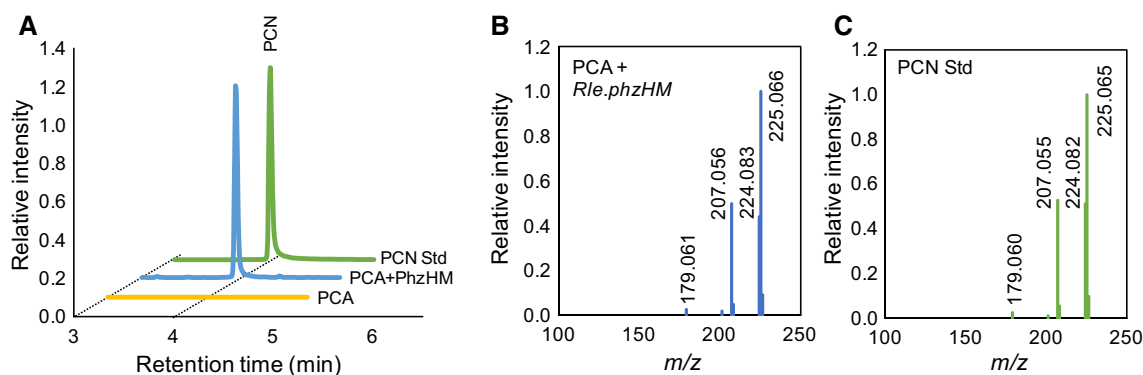
Fig. 6 Production of PCN. PCN production of *E. coli* BL21(DE3) harboring pW5PCA-*Rle.phzHM* is shown. The LC–MS chromatogram of *m/z* 224.08 (**a**), the MS2 spectra of *m/z* 224.08 [M+H]⁺ from the culture extract of *E. coli* with the core phenazine BGC and *phzHM* (**b**) and PCN chemical standard (**c**). Although it is unclear why MS2 spectra showed *m/z* 225.065, these two MS2 spectra superpose each other. In these panels, the culture extracts of *E.* *coli* BL21(DE3) harboring pW5PCA and pW5PCA-*Rle.phzHM* are shown in light orange and blue, respectively. The PCN chemical standard is shown in green. No peak corresponding to 5-methyl-PCN (*m/z* 238.08) was seen. MS2 collected at 30 eV collision energy. Each LC–MS trace/MS2 spectrum is a representative of triplicate experiments (color figure online)
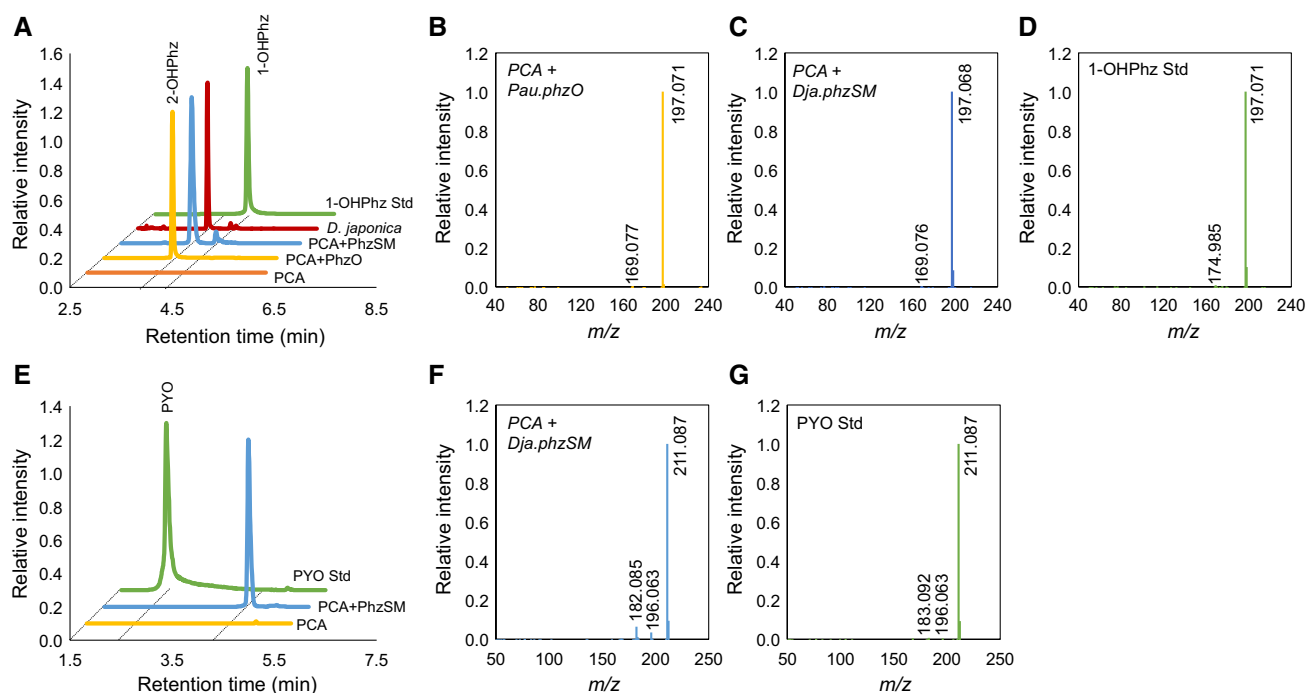

Fig. 7 Production of 2-OHPhz and a PYO isomer. 2-OHPhz and PYO production of *E. coli* BL21(DE3) harboring pW5PCA-*Dja.phzSM* is shown. The LC–MS chromatogram of *m/z* 197.07 (**a**), the MS2 spectra of peaks with *m/z* 197.07 [M+H]⁺ of the culture extracts of *E. coli* with the core phenazine BGC, native *D. japonica* and *Pau.phzO* (**b**), *phzSM* (**c**), and the 1-OHPhz chemical standard (**d**), the LC–MS chromatogram of *m/z* 211.09 (**e**), and the MS2 spectra of peaks with *m/z* 211.09 [M+H]⁺ of the culture extract of *E. coli* with the core phenazine BGC and *phzSM* (**f**) and the PYO chemical standard (**g**). In these panels, the culture extracts of *E. coli* BL21(DE3) harboring pW5PCA and pW5PCA-*Pau.phzO*, and pW5PCA-*Dja.phzSM* are shown in dark orange, light orange, and blue, respectively. The 1-OHPhz and PYO chemical standards are shown in green. MS2 collected at 30 eV collision energy. Each LC–MS trace/MS2 spectrum is a representative of triplicate experiments (color figure online)

from PhzM alone would be 5-methyl-PCA; however, we were also unable to detect this molecule. The *D. japonica* module consists of *phzS* and *phzM* homologs. Although we expected to see 1-OHPhz and PYO, we detected 2-OHPhz and an isomer of PYO as notable products (Figs. 2, 7). LC–MS was used to identify the retention time and accurate mass for the 1-OHPhz authentic standard (Fig. 7a). Because 2-OHPhz is not commercially available, we cloned
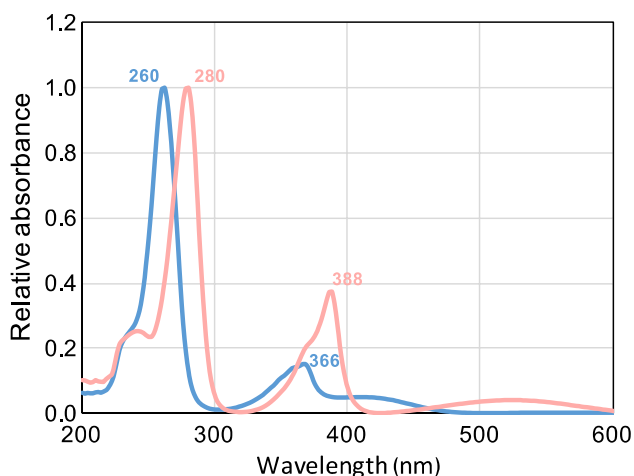
Fig. 8 Background subtracted UV–Vis spectra for PYO and a PYO isomer. UV–Vis spectra for PYO (pink) and the PYO isomer (blue) are shown. These spectra were obtained during LC–MS run, and background spectra for their corresponding negative controls were subtracted (color figure online)

and over-expressed *phzO* from *P. aureofaciens* 30–84 along with the PCA-producing core phenazine BGC in *E. coli* to prepare 2-OHPhz. PhzO was previously characterized to catalyze reductive hydroxylation at the 2-position [5]. The notable OHPhz product matched the LC retention time, exact mass, and MS2 of the high-resolution 2-OHPhz peak (Fig. 7a–c). The *D. japonica* native extract also exhibited a peak that matches the exact mass and retention time of 2-OHPhz (Fig. 7a). From this, we surmise that the *phzS* gene homolog from *D. japonica* was responsible for hydroxylation at primarily the 2-position rather than the 1-position (Figs. 2, 7a–d). The comparison of the PYO isomer to the pyocyanin chemical standard showed a different retention time and MS2 spectra, while both molecular ions shared the identical mass (Fig. 7e–g). UV–Vis analyses for PYO and the PYO isomer showed similar, but distinctive spectra (Fig. 8).

Discussion

Functional characterization of secondary metabolite BGCs is an essential step toward understanding the roles of the secondary metabolites in complex ecosystems [14]. However, these BGCs are often very difficult to characterize, as their function is highly regulated with complex mechanisms. Synthetic biology offers promising tools and methods to completely rewrite genetic codes, which may allow elimination of complex regulation embedded in the native BGCs. Therefore, we developed an integrated workflow of BGC mining, refactoring, and DNA synthesis and assembly and investigated its efficacy using three phenazine BGCs,

selected from bioinformatics analysis through the IMG/ABC database, as models. We decided to first characterize the function of five genes from three BCGs homologous to known PCA-modifying enzymes from these BGCs (including the *phzO* gene). These genes were over-expressed along with the core Phz BGCs.

Rhizobium leguminosarum is an α-proteobacterium found in a clover nodule system and likely plays an important role in nitrogen fixation. We observed only PhzH activity in *E. coli* expressing both *phzH* and *phzM* homologs from *R. leguminosarum*. Demonstration of PCN production by this construct is significant, because the Phz BGC found in *R. leguminosarum* is the first instance of Phz production in α-proteobacteria. We also confirmed PCA production in the culture extract of *R. leguminosarum* (Fig. 5d, e). Despite successful cloning and likely expression of PhzM, none of the predicted 5- or 10-methyl PCN products were detected. PhzM may not be expressed at sufficient levels or the PhzM homolog in the *R. leguminosarum* pathway may require a different substrate than the PCA or PCN produced. The four remaining genes in this BGC, including *phzM,* may be involved in decoration of PCN. To the best of our knowledge, phenazines derived from PCN have not been previously reported. Further characterization of this unique BGC may reveal if the phenazine produced from this pathway plays a role in the clover root nodule system such as respiration mechanisms of *R. leguminosarum* in anaerobic/semi-aerobic conditions. Such mechanisms may also be important in oxygen-sensitive nitrogen fixation enzymes and/or in controlling the population of other microbes within the nodule systems itself.

Dyella japonica is a γ-proteobacterium isolated from an *Arabidopsis* root system. Our studies show that over-expression of *phzSM* homologs from *D. japonica* were able to modify PCA to form 2-OHPhz and the PYO isomer in *E. coli*. To our surprise, the PhzS homolog from *D. japonica* mainly catalyzed the formation of 2-OHPhz from PCA as a substrate. The same activity was previously identified in *P. aureofaciens* 30–84, and the *phzO* gene was identified as responsible for this activity. PhzM is known to catalyze N5-methylation of 1-OHPhz to form PYO. However, this PhzM homolog produced the PYO isomer. Because the PhzS homolog produced mainly 2-OHPhz, the PhzM homolog likely catalyzed N-methylation of 2-OHPhz as a substrate.

Production of the PYO isomer was confirmed with comparison of the retention time, high-resolution mass, MS2, UV–Vis spectra for those of the PYO standard. There are only four possible isomeric structures for PYO. Among them, two PYO isomers, 5-methyl-1-OHPhz (PYO) and 10-methyl-2-OHPhz (10-Me-2-OHPhz), are likely more stable than the other structures, because PYO and 10-Me-2-OHPhz can form deprotonated states (Fig. 9). Thus, it is speculated that the observed molecule is likely
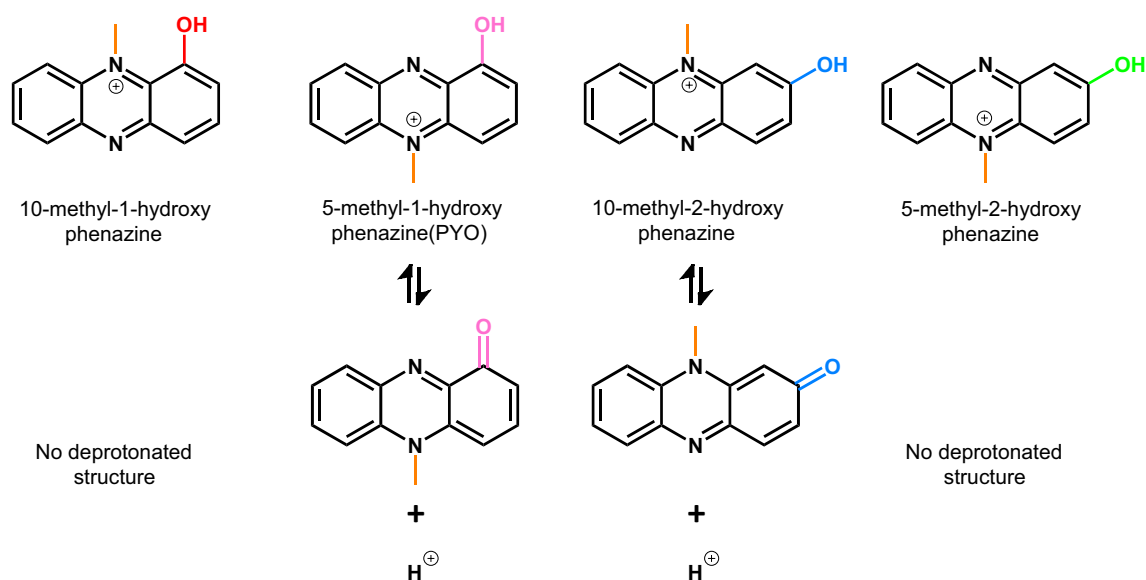
Fig. 9 Possible PYO structural variants considered for identification of the compound produced by *E. coli* expressing PCA-*Dja.phzSM*. Protonated and deprotonated structural variants are displayed above and below. 10-Methyl-1-hydroxy phenazine and 5-methyl-2-hydroxy phenazine are less stable because deprotonation of these compounds does not result in a stable structure while 5-methyl-1-hydroxy (PYO) and 10-methyl-2-hydroxy can be deprotonated

10-Me-2-OHPhz. Although we will need further validation using NMR, to the best of our knowledge, this is the first time that a biological route for production of the PYO isomer has been reported [27, 28].

Our characterization of four phenazine-modifying enzymes provide additional evidence that the integrated workflow that we employed is useful to simplify and accelerate functional characterization of BGCs and set the blueprint for going after novel chemical diversity. We successfully characterized three new phenazine-modifying genes from previously uncharacterized phenazine BGCs from unique plant-associated microbes. Although we primarily focused on characterizing homologs of known Phz-modifying genes in this study, because of the modular nature of the Phz BGCs, the same approach is currently being utilized to validate the function of remaining genes in these BGCs, which is expected to further prove the utility of the integrated workflow. Although phenazine derivatives are known for their antibiotic activity, we did not observe changes in growth of the production *E. coli* strain (at least with our production level). However, use of *Pseudomonas* spp. as production hosts may be more useful for characterization of new Phz BGCs to minimize unintended toxicity issues. Additionally, because there is also a possibility that heterologous expression in *E. coli* yields different products from native ones, use of native PCA producers such as *P. synxantha* 2-79 [10] as hosts may be also useful for confirming function of three new phenazine-modifying enzymes described here and characterizing function of remaining genes in these BGCs.

Finally, an advantage of this integrated workflow is the automated nature of module refactoring using the BOOST tools that can fully automate codon juggling to fragmentation. This automated workflow for refactoring can be applied to any pathway. The modular approach to pathway characterization is somewhat limited to BCGs that exhibit a core pathway with a diversity of modifying enzymes. Uncharacterized pathways may still require cloning of the native sequence or synthesis and refactoring of pathways to gain sufficient information needed to expand to modular characterization. However; for characterized pathways, the combination of DNA synthesis of modifying enzymes and recombineering with a core pathway can be a powerful strategy to test a broad diversity of combinations to explore new chemical diversity without the cost of repeated synthesis or cloning of the entire sequence.

References

1. Blankenfeldt W, Parsons JF (2014) The structural biology of phenazine biosynthesis. Curr Opin Struct Biol 29:26–33. https://doi.org/10.1016/j.sbi.2014.08.013

2. Chen et al. 2015. DIVA: more science, less DNA construction. 2015 Synth Biol Eng Evol Des (SEED) Proc Poster

3. Chin-A-Woeng TFC, Thomas-Oates JE, Lugtenberg BJJ, Bloemberg BJJ (2001) Introduction of the *phzH* gene of *Pseudomonas chlororaphis* PCL1391 extends the range of biocontrol ability of phenazine-1-carboxylic acid-producing *Pseudomonas* spp. strains. MPMI 14(8):1006–1015. https://doi.org/10.1094/mpmi.2001.14.8.1006

4. Datsenko KA, Wanner BL (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc Natl Acad Sci USA 97(12):6640–6645. https://doi.org/10.1073/pnas.120163297

5. Delaney SM, Mavrodi DV, Bonsall RF, Thomashow LS (2001) *phzO*, a gene for biosynthesis of 2-hydroxylated phenazine compounds in *Pseudomonas aureofaciens* 30-84. J Bacteriol 183(1):318–327. https://doi.org/10.1128/JB.183.1.318-327.2001

6. Kushwaha M, Salis HM (2015) A portable expression resource for engineering cross-species genetic circuits and pathways. Nat Commun 6:7832. https://doi.org/10.1038/ncomms8832

7. Hadjithomas M, Chen IMA, Chu K et al (2015) IMG-ABC: a knowledge base to fuel discovery of biosynthetic gene clusters and novel secondary metabolites. MBio 6(4):e00932-15. https://doi.org/10.1128/mBio.00932-15

8. Hillson JN, Rosengarten RD, Keasling JD (2012) j5 DNA assembly design automation software. ACS Synth Biol 1(1):14–21. https://doi.org/10.1021/sb2000116

9. Mavrodi DV, Ksenzenko VN, Bonsall RF, Cook RJ, Boronin AM, Thomashow LS (1998) A seven-gene locus for synthesis of phenazine-1-carboxylic acid by *Pseudomonas fluorescens* 2-79. J Bacteriol 180(9):2541–2548

10. Mavrodi DV, Bonsall RF, Delaney SM, Soule MJ, Phillips G, Thomashow LS (2001) Functional analysis of genes for biosynthesis of pyocyanin and phenazine-1-carboxamide from *Pseudomonas aeruginosa* PAO1. J Bacteriol 183:6454–6465. https://doi.org/10.1128/JB.183.21.6454-6465.2001

11. Mavrodi DV, Blankenfeldt W, Thomashow LS (2006) Phenazine compounds in fluorescent *Pseudomonas* spp. biosynthesis and regulation. Annu Rev Phytopathol 44:417–445. https://doi.org/10.1146/annurev.phyto.44.013106.145710

12. Mavrodi DV, Peever TL, Mavrodi OV et al (2010) Diversity and evolution of the phenazine biosynthesis pathway. Appl Environ Microbiol 76(3):866–879. https://doi.org/10.1128/AEM.02009-09

13. Oberortner E, Cheng JF, Hillson NJ, Deutsch S (2017) Streamlining the design-to-build transition with build-optimization software tools. ACS Synth Biol 6(3):485–496. https://doi.org/10.1021/acssynbio.6b00200

14. Owen JG, Reddy BVB, Ternei MA et al (2013) Mapping gene clusters within arrayed metagenomic libraries to expand the structural diversity of biomedically relevant natural products. Proc Natl Acad Sci 110(29):11797–11802. https://doi.org/10.1073/pnas.1222159110

15. Pierson LS, Thomashow LS (1992) Cloning and heterologous expression of the phenazine biosynthetic locus from *Pseudomonas aureofaciens* 30-84. Mol Plant Microbe Interact 5:330–339. https://doi.org/10.1094/MPMI-5-330

16. Pierson EA, Wang D, Pierson LS III (2013) Roles and regulation of phenazines in the biological control strain *Pseudomonas chlororaphis* 30-84. In: Chincholkar S, Thomashow LS (eds) Microbial phenazines. Springer, Berlin, pp 141–162

17. Glasser NR, Kern SE, Newmann DK (2007) Phenazine redox cycling enhances anaerobic survival in *Pseudomonas aeruginosa* by facilitating generation of ATP and a proton-motive force. Mol Microbiol 92(2):1365–2958. https://doi.org/10.1111/mmi.12566

18. Ren H, Wang B, Zhao H (2017) Breaking the silence: new strategies for discovering novel natural products. Curr Opin Biotechnol 48:21–27. https://doi.org/10.1016/j.copbio.2017.02.008

19. Rui Z, Ye M, Wang S et al (2012) Insights into a divergent phenazine biosynthetic pathway governed by a plasmid-born esmeraldin gene cluster. Chem Biol 19:1116–1125. https://doi.org/10.1016/j.chembiol.2012.07.025

20. Smanski MJ, Zhaou H, Claesen J, Shen B, Fischbach MA, Voigt CA (2016) Synthetic biology to access and expand nature's chemical diversity. Nat Rev Microbiol 14(3):135–149. https://doi.org/10.1038/nrmicro.2015.24

21. Santos CNS, Yoshikuni Y (2014) Engineering complex biological systems in bacteria through recombinase-assisted genome engineering. Nat Protoc 9(6):1320–1336. https://doi.org/10.1038/nprot.2014.084

22. Thomashow LS, Bonsall RF, Weller DM (2002) Antibiotic production by soil and rhizosphere microbes in situ. In: Hurst CJ (ed) Manual of environmental microbiology. ASM Press, Washington, pp 638–647

23. Thomashow LS (2013) Phenazines in the environment: microbes, habitats, and ecological relevance. In: Chincholkar S, Thomashow LS (eds) Microbial phenazines. Springer, Berlin, pp 199–216

24. Cheng X, Zhang X, Pflugrath JW, Studier FW (1994) The structure of bacteriophage T7 lysozyme, a zinc amidase and an inhibitor of T7 RNA polymerase. Proc Natl Acad Sci USA 91(9):4034–4038

25. Ikeda RA (1992) The efficiency of promoter clearance distinguishes T7 class II and class III promoters. J Biol Chem 267(16):11322–11328

26. Gibson GG, Young L, Chuang RY, Venter JC, Hutchison CA III, Smith HO (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat Methods 6:343–345. https://doi.org/10.1038/nmeth.1318

27. Kehmann F, Cherpillod F (1924) Neue Synthesen in der Gruppe der Chinon-imid-farbstoffe. V. Über Synthesen ausgehend von Oxybenzochinon. Helv Chim Acta 7(1):973–980. https://doi.org/10.1002/hlca.192400701122

28. Mcllwain H (1937) The phenazine series. Part VI. Reactions of alkyl phenazonium salts; the phenazyls. J Chem Soc 0:1704–1711. https://doi.org/10.1039/JR9370001704