Check for updates

Proteomining-Based Elucidation of Natural Product Biosynthetic Pathways in *Streptomyces*

Darwin Linardi[1†], Weiyi She[2,3,4†], Qian Zhang[5], Yi Yu[5], Pei-Yuan Qian[2,4] and Henry Lam[1]*

[1] Department of Chemical and Biological Engineering, The Hong Kong University of Science and Technology, Kowloon, Hong Kong SAR, China, [2] Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou), Hong Kong, Hong Kong SAR, China, [3] SZU-HKUST Joint PhD Program in Marine Environmental Science, Shenzhen University, Shenzhen, China, [4] Department of Ocean Science and Hong Kong Branch of Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou), The Hong Kong University of Science and Technology, Kowloon, Hong Kong SAR, China, [5] Hubei Clinical Center and Key Laboratory of Intestinal and Colorectal Disease, Department of Gastroenterology, Zhongnan Hospital of Wuhan University, School of Pharmaceutical Sciences, Wuhan University, Wuhan, China

The genus *Streptomyces* is known to harbor numerous biosynthetic gene clusters (BGCs) of potential utility in synthetic biology applications. However, it is often difficult to link uncharacterized BGCs with the secondary metabolites they produce. Proteomining refers to the strategy of identifying active BGCs by correlating changes in protein expression with the production of secondary metabolites of interest. In this study, we devised a shotgun proteomics-based workflow to identify active BGCs during fermentation when a variety of compounds are being produced. Mycelia harvested during the non-producing growth phase served as the background. Proteins that were differentially expressed were clustered based on the proximity of the genes in the genome to highlight active BGCs systematically from label-free quantitative proteomics data. Our software tool is easy-to-use and requires only 1 point of comparison where natural product biosynthesis was significantly different. We tested our proteomining clustering method on three *Streptomyces* species producing different compounds. In *Streptomyces coelicolor* A3(2), we detected the BGCs of calcium-dependent antibiotic, actinorhodin, undecylprodigiosin, and coelimycin P1. In *Streptomyces chrestomyceticus* BCC24770, 7 BGCs were identified. Among them, we independently re-discovered the type II PKS for albofungin production previously identified by genome mining and tedious heterologous expression experiments. In *Streptomyces tenebrarius*, 5 BGCs were detected, including the known apramycin and tobramycin BGC as well as a newly discovered caerulomycin A BGC in this species. The production of caerulomycin A was confirmed by LC-MS and the inactivation of the caerulomycin A BGC surprisingly had a significant impact on the secondary metabolite regulation of *S. tenebrarius*. In conclusion, we developed an unbiased, high throughput proteomics-based method to complement genome mining methods for the identification of biosynthetic pathways in *Streptomyces* sp.

Keywords: *Streptomyces*, proteomics, synthetic biology, natural product, *Actinobacteria*, biosynthetic gene cluster (BGC), antibiotic

INTRODUCTION

The *Streptomyces* sp. are producers of natural products that are often synthesized by genes in proximity in the genome, collectively referred to as biosynthetic gene clusters (BGCs) (Hopwood, 2007; Medema et al., 2015; Genilloud, 2017). Many of these BGCs share common core genes such as polyketide synthases (PKS) and non-ribosomal peptide synthetases (NRPS), which form an assembly line that elongates the metabolite backbone. The combination of these core genes with various functional groups and structure modifying enzymes in a BGC allows the *Streptomyces* sp. to produce a diverse variety of compounds that are difficult to efficiently produce through chemical synthesis (Hopwood, 2007; Rossiter et al., 2017; Genilloud, 2018).

The democratization of genome sequencing has revealed the variety of BGCs in the genomes of *Streptomyces* spp., which contributes to the recent interest in leveraging this genus for synthetic biology applications (Genilloud, 2018). This led to the increased popularity of genome mining, the identification of novel BGCs in newly sequenced genomes via sequence analysis paired with experimental confirmation of the products of the BGCs. One of the most widely used genome mining tools is antiSMASH, which predicts BGCs in the genome and matches the predicted clusters against a database of previously documented BGCs, allowing for rapid identification of putative BGCs in a newly sequenced genome (Blin et al., 2021). Another important genome mining tool is PRISM which uses machine learning to predict the products produced by the identified BGCs (Skinnider et al., 2015). However, putative BGCs predicted by genome mining are not always specific, and many are not actively expressed or "silent." In addition, since many BGCs are present in a single genome of *Streptomyces* sp., linking predicted BGCs to their respective metabolites can be challenging.

Complementing genome mining approaches with expression-based analyses, such as transcriptomics or proteomics, can help to focus the search for the BGC(s) responsible for the production of compound(s) of interest by correlating secondary metabolite production levels to gene expression levels (Gubbens et al., 2014). For example, transcriptomics has been performed on *Streptomyces flaveolus* (Qu et al., 2011) and on several strains of *Salinispora pacifica* (Amos et al., 2017) to search for BGCs. However, transcripts have short half-lives in bacteria, making it difficult to determine the optimal time points to capture the shift in expression patterns (Moran et al., 2013). On the contrary, proteins are the actual actors in natural product biosynthesis in real-time; thus, one expects that protein levels should be more reliably correlated with secondary metabolite production (Du and van Wezel, 2018; Ludvigsen and Honoré, 2018). The term "proteomining" was coined by Gubbens et al. (2014) to describe the strategy of identifying potential BGCs by correlating changes in the protein expression pattern, measured by proteomics, with the production of secondary metabolites of interest.

The early proteomining methods targeted the phosphopantetheinyl cofactor of PKS and NRPS (Meluzzi et al., 2008; Bumpus et al., 2009). They utilized the mass shift that occurs due to the ionization-induced release of 4′-phosphopantetheine to identify spectra corresponding to PKS/NRPS peptides linked to this cofactor. The sequence of these peptides was then determined *de novo* and reverse translated to develop primers for the targeted PCR study of the gene. To improve the selectivity for peptides of PKS/NRPS, Meier et al. (2009, 2011) developed probes that bind to the conserved active sites or the phosphopantetheinyl cofactor to selectively pull down expressed PKS and NRPS. In another approach, high molecular weight proteins were isolated from streptomycetes to select for PKS and NRPS complexes (Bumpus et al., 2009; Evans et al., 2011; Chen et al., 2012, 2013a,b). However, these approaches are "biased," being able to detect only NRPS and PKS, and do not cover the whole proteome but have the advantage of not requiring the whole genome sequence. To this end, an unbiased method employing shotgun proteomics to compare the proteomes of mycelia grown on different media was proposed (Gubbens et al., 2014). Using this method, the authors linked the expressed BGCs to antibiotic activity and subsequently identified the biosynthetic pathway of a novel juglomycin antibiotic.

In this work, we exploited the actinomycetes' characteristic of secondary metabolite BGCs expression during nutrient starvation and developed a label-free proteomics-based method for proteomining. The non-producing, nutrient-rich growth phase was used as background to the secondary metabolite producing, nutrient-starved stationary phase. This served as another axis of comparison alongside comparison between mutant strains or different fermentation media. The contrast between the producing/non-producing proteome states, combined with the clustering of differentially expressed proteins based on proximity in the genome, would pinpoint expressed BGCs. To demonstrate the utility of the method, three *Streptomyces* sp. were studied in this work. First, we applied the method to the well-studied *Streptomyces coelicolor* A3(2) as a proof of concept. Then, we sought to identify active BGCs in *Streptomyces chrestomyceticus* BCC24770 and *Streptomyces tenebrarius*.

EXPERIMENTAL METHODS

Strain Cultivation Conditions

Fermentation of *S. coelicolor* A3(2) and *S. chrestomyceticus* BCC24770 was performed at 28°C and 250 RPM with 30 glass beads per flask. Fermentation of *S. tenebrarius* was performed at 37°C and 250 RPM. Spores of *S. coelicolor* A3(2) and BCC24770 were activated in GYM broth (4 g glucose, 4 g yeast extract, and 10 g malt extract per liter Milli-Q water, DSMZ Medium 65) for 24 to 48 h (OD_{600} = 1) and transferred to fresh GYM broth (1% inoculum) to start the fermentation. Fermentation of BCC24770 in TSBY broth (30 g tryptic soy broth and 5 g yeast extract per liter of Milli-Q water) media was inoculated using GYM broth culture at OD_{600} = 1 (1% inoculum). Similarly, spores of *S. tenebrarius* were activated in TSBY for 24 h and transferred to fresh TSBY broth. Fermentation of *S. tenebrarius* was performed at 37°C and 250 RPM. All fermentation experiment was performed for up to 3 days. The harvested fermentation liquid was separated

by centrifugation and both fractions were stored at −20°C until sample processing.

Secondary Metabolite Characterization

Undecylprodigiosin was extracted and characterized based on the method previously developed (Tenconi et al., 2013). In summary, the mycelia of *S. coelicolor* A3(2) at 96 h were incubated in acidified methanol (pH < 4) for 1 h. The suspension was centrifuged to remove cell debris. The absorbance spectra of the supernatant were measured using Varioskan™ LUX multimode microplate reader. Undecylprodigiosin-hydrochloride (Sigma Aldrich) was used as standard. Mass spectrometry analysis was performed by direct injection to a SCIEX TripleTOF 4600.

Apramycin concentration in the supernatant was determined by a Waters Acquity HPLC coupled with an evaporative light scattering detector (ELSD) (Lv et al., 2016; Zhang et al., 2021). The column used was Waters CORTECS™ UPLC C18 1.6 μm (2.1 × 100 mm). Mobile phases A and B were 20 mM heptafluorobutyric acid in water and acetonitrile, respectively. The flow rate of the HPLC was set at 0.3 mL/min. The gradient used started with 10% B, increasing linearly to 50% B over 10 min, then to 100% B in 2 min. The mobile phase is held at 100% B for 3 min to purge the system before returning to 10% B for 5 min in preparation for the next injection. The ELSD was set at $65 \pm 25°C$ and 40.0 psi gas pressure. Apramycin sulfate (Sigma Aldrich) was used as standard for peak identification.

The extraction of caerulomycin A from the fermentation of *S. tenebrarius* was based on a method described previously (Zhu et al., 2012). The supernatant was extracted using equal volume ethyl acetate and vigorously shaken. The process was repeated up to three times. The organic fraction was pooled and concentrated using a vacuum concentrator. The residual solid was redissolved in methanol and analyzed by HPLC-MS. The column used was Waters Acquity UPLC BEH C18 Column 1.7 μm (2.1 × 150 mm). The gradient used was the same as used in apramycin quantification, except that the mobile phases were 1% formic acid in water and acetonitrile as A and B, respectively.

The crude extract of BCC24770 was obtained as described previously (She et al., 2020, 2021). Briefly, an equal volume of ethyl acetate was added to the fermentation supernatant and vigorously shaken for 3 h. The mixture was centrifuged at 14,000 *g* for 10 min. The organic phase was pooled in a fresh container. This process was repeated three times in total. Ethyl acetate was evaporated using a vacuum concentrator and the remaining solid was resuspended in methanol and analyzed by HPLC-UV/Vis. The column used was Waters CORTECS™ UPLC C18 1.6 μm (2.1 × 100 mm). The gradient used was the same as used in apramycin quantification, except that the mobile phases were 0.1% trifluoroacetic acid in water and acetonitrile as A and B, respectively.

Shotgun Proteomics

Shotgun proteomics sample preparation was prepared by a modified acetone precipitation protocol followed by in-solution digestion (Liu et al., 2016; Hao and Lam, 2021). Mycelia were resuspended in lysis buffer (0.5% sodium dodecyl sulfate, 50 mM Tris–HCl, pH 8.5) and incubated at 95°C for 5 min. The suspension was sonicated in an ice bath until translucent and centrifuged at 14,000 *g* for 10 min to remove cell debris. Ice cold acetone was added to the supernatant at a 4:1 ratio and incubated at −20°C overnight. Purified protein was pelleted by centrifugation at 14,000 *g* and 0°C for 30 min and further washed with ice-cold washing solution (80% acetone, 10% methanol, 0.2% acetic acid) to remove sodium dodecyl sulfate attached to the proteins. The protein pellets were then resuspended in 6 M urea, 0.6 M guanidine hydrochloride, and 50 mM ammonium bicarbonate and quantified by bicinchoninic acid assay (Pierce™, Thermo Fisher). A 40 μg aliquots of the protein solution were reduced with 0.1 M dithiothreitol (Sigma) at 37°C for 1 h. Exposed cysteine residues were methylated by 0.2 M iodoacetamide (Sigma) in the dark for 30 min, after which the reaction was quenched using an excess of dithiothreitol. The protein solution was diluted with 50 mM ammonium bicarbonate to a final concentration of 1 M urea. Sequencing-grade modified trypsin (Promega) was used at a 1:50 trypsin to protein ratio. Trypsin digestion was performed overnight at 37°C. Formic acid was added to a final concentration of 0.5% to quench the digestion. The peptide mixture was desalted using C18 ZipTip (Merck Millipore) as per the manufacturer's instruction. Desalted peptides were stored at −20°C until analysis. Peptides were resuspended in 1% formic acid at 200 ng/μL and injected into Bruker timsTOF Pro Mass-Spectrometer. LCMS operation parameter is listed in **Supplementary Table 1**.

The LCMS raw data was converted by the msconvert.exe tool of the ProteoWizard suite (Ver. 3.0.21101) (Chambers et al., 2012) to mzXML with the following filters: top 150 most-intense count, 0.05 m/z precursor tolerance, 5 s scan time tolerance, and 0.1 ion mobility tolerance. The *S. coelicolor* A3(2) and BCC24770 protein database used consisted of open reading frames (ORFs) generated by GeneMarkS (Besemer et al., 2001). The *S. tenebrarius* protein database was obtained from the RAST database (Aziz et al., 2008). The genome of *S. coelicolor* A3(2) was obtained from the NCBI Genome database (Bentley et al., 2002). The genome of BCC24770 was sequenced and deposited in a previous study (She et al., 2021). Protein annotation for the *S. coelicolor* A3(2) protein database was performed using BLASTP of the NCBI Blast Suite against *Actinobacteria* proteins downloaded from UniProtKB on 11th November 2021 at 1E-3 *E*-value cutoff (Camacho et al., 2009). Protein identification was performed using the *Trans*-Proteomic Pipeline (Deutsch et al., 2011; Ver, 2021). Peptide-spectrum matches were obtained by sequence database searching using Comet (Ver, 2021). The settings used were precursor peptide mass tolerance of 30 ppm and fragment ion bin tolerance of 0.05. The peptide-spectrum matches were validated by PeptideProphet, iProphet, and ProteinProphet with the false detection rate set at 1%. Label-free quantification was performed by calculating the non-spectral abundance factor (NSAF) (McIlwain et al., 2012) when the average spectral count in a sample group ≥ 3. Proteins that did not pass this threshold or were not detected in the sample were assigned the value of 1/10 of the lowest NSAF.

Biosynthetic Gene Cluster Identification Algorithm

The fold-change and the *p*-value of pairwise comparisons were calculated using the average NSAF values of each sample group. Differentially expressed proteins were assigned as nodes that scored 10 points when the node was highly differentially expressed (magnitude of the fold-change ≥ 2.0 and *p*-value ≤ 0.01) or five points when the node was moderately differentially expressed (magnitude of the fold-change ≥ 1.5 and *p*-value ≤ 0.05). A positive score was assigned to ORFs that were upregulated at the nutrient-starved phase and a negative score was assigned to ORFs that were expressed during the growth phase. Edges were drawn between nodes that are within 5 ORFs of each other. Each cluster was scored based on the magnitude

of the average score of all the nodes. Clusters with scores ≥ 3 are deemed a putative cluster of differentially expressed proteins, hereafter referred to as "proteomining-based clusters" (PBCs) to distinguish from BGCs defined by genome mining. Those with scores between 1.5 and 3 were checked for chimeric clusters by manual inspection. If there exists a partition of the cluster whereby the sub-clusters have score ≥ 3, the chimeric cluster is separated, and the sub-clusters are annotated as a PBC with a letter. These score cutoffs were chosen based on experience with our data. The remaining PBCs containing at least 5 nodes and scored greater than 3 were annotated. Functional analysis of the individual PBCs formed in the proteomics analysis of *S. coelicolor* A3(2) was performed using STRING (Ver. 11.5) with "*Streptomyces coelicolor* A3(2)" as the organism, the
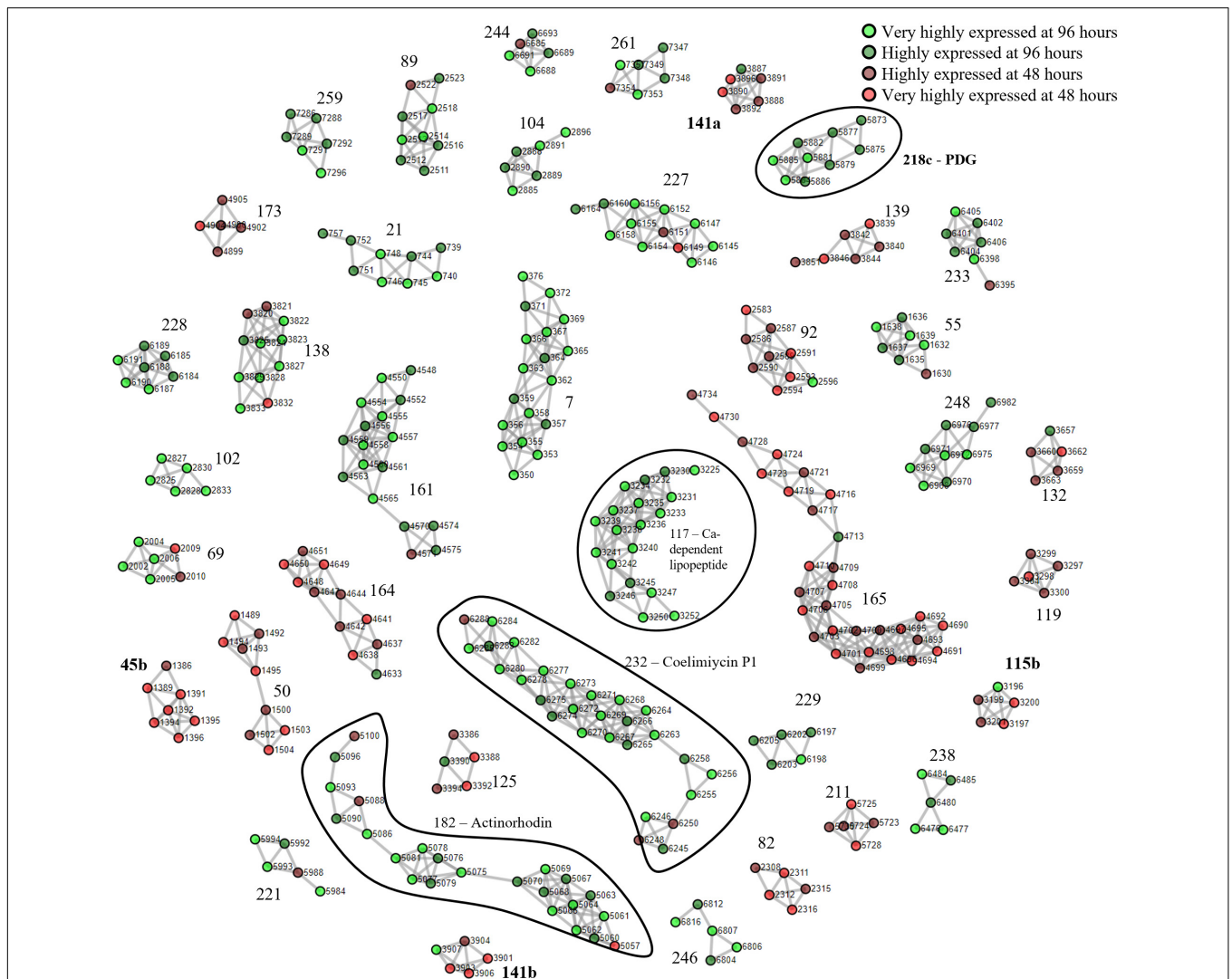

FIGURE 1 | Proteomining result of *S. coelicolor* A3(2) fermentation in GYM medium. ◉ and ◉ represent proteins that were more abundant in the mycelia harvested at 96 h compared to 48 h at foldchange ≥ 2 and *p*-value ≤ 0.01 (very highly expressed at 96 h), and foldchange ≥ 1.5 and *p*-value ≤ 0.05 (highly expressed at 96 h), respectively. ◉ and ◉ represent proteins that were more abundant in the mycelia harvested at 48 h compared to 96 h at foldchange ≥ 2 and *p*-value ≤ 0.01 (very highly expressed at 48 h), and foldchange ≥ 1.5 and *p*-value ≤ 0.05 (highly expressed at 48 h), respectively. Edges were drawn between nodes with an ORF difference ≤ 5. Cluster numbers in bold indicate clusters that previously scored between 1.5 and 3 which scored ≥ 3 when the chimeric clusters were separated.

"Neighborhood" component of active interaction score disabled, and minimum interaction score set at 0.4 (Medium) (Szklarczyk et al., 2021). The STRING cluster node functions were checked against the STRING database for overall PBC function, if available. BGC characterization was performed by antiSMASH Ver 6.0 using the default settings (Blin et al., 2021). The in-house scripts implementing the proteomining clustering algorithm are available at https://github.com/darwinlinardi/Proteomining-Clustering.

Streptomyces tenebrarius ORF 4229–4230 Knockout

DNA fragments 2,486 bp upstream and 2,436 bp downstream of ORF 4229-4230 (Gene name: *crmA*) was amplified by PCR using the *S. tenebrarius* genome as template. Purified fragments were cloned into the *Hind*III/*Eco*RI sites of pWHU258 (Lv et al., 2016) gene knockout vector using Hieff Clone® Plus Multi One Step Cloning Kit (Yeasen, China). The vector was inserted into *S. tenebrarius* via *E. coli-Streptomyces* conjugation (Kieser et al., 2000). The resulting in-frame deletion mutants *S. tenebrarius* Δ*crmA* were screened and confirmed by PCR analysis and DNA sequencing. The primers used in the knockout are displayed in **Supplementary Table 2**.

RESULTS AND DISCUSSION

Streptomyces coelicolor A3(2) Biosynthetic Gene Cluster Proteomining

We first used *S. coelicolor* A3(2) as a benchmark due to its well-documented expression of BGCs that allowed for clear identification of the clusters. The mycelia harvested at 96 h were designated as the nutrient-starved phase which was indicated by the red pigmentation. The mycelia harvested at 48 h were designated as the growth phase. The described clustering algorithm was applied to yield putative PBCs. The devised PBC scoring system rewarded consistent upregulation of cluster members in the same sample group, divided by the difference between the smallest and largest ORF numbers, to encourage the formation of tight clusters. A total of 57 PBCs had a score $\geq$ 1.5 and among these, 35 had a score $\geq$ 3 (**Supplementary Figure 1** and **Supplementary Data 1**). On examination of the cluster structure, 4 of the 22 PBCs with scores between 1.5 and 3 (PBC No. 45, 115, 141, and 218) were decoupled and subsequently formed sub-clusters that scored $\geq$ 3. In total, 40 PBCs were identified and are displayed in **Figure 1**. The function of the PBCs was predicted using STRING (Ver 11.5) with the "neighborhood" criterion disabled to increase orthogonality between our proteomining clustering method and STRING since our method had already determined that the proteins in each PBC were already in proximity in the genome (**Table 1**).

During the nutrient starvation phase, we detected the upregulation of PBC No. 117, 182, 218c, and 232, annotated by antiSMASH as the calcium-dependent antibiotic, actinorhodin, undecylprodigiosin, and coelimycin P1 BGCs, respectively. The production of calcium-dependent antibiotic, actinorhodin, and undecylprodigiosin under nutrient starvation conditions has been well-documented even before the whole-genome assembly of *S. coelicolor* A3(2) in 2003 (Bentley et al., 2002; Hopwood, 2007; Manteca et al., 2010; Bednarz et al., 2019). However, PBC No. 218c, corresponding to the undecylprodigiosin BGC, required decoupling for the PBC to score $\geq$ 3. To ensure that

TABLE 1 | Functional annotation of proteomining clusters of *S. coelicolor* A3(2) at 48 and 96 h.

Cluster	ORF begin	ORF end	Members	Highly expressed at	Cluster function
7	350	376	18	96 h	Polysaccharide biosynthesis
50	1489	1504	9	48 h	Ribosomal proteins
55	1630	1639	7	96 h	Proteasomes
69	2002	2010	6	96 h	Amino acid transport
82	2308	2316	5	48 h	Protein export
89	2511	2523	9	96 h	Sensor
92	2583	2596	9	48 h	Ribosomal proteins
102	2825	2833	5	96 h	Maleate isomerase
115b	3196	3201	5	48 h	Fructose/mannose metabolism
117	**3225**	**3252**	**19**	**96 h**	**Ca-dependent lipopeptide biosynthesis**
132	3657	3663	5	48 h	Stress response
141b	3901	3907	5	48 h	Ribosomal proteins
161	4548	4575	17	96 h	NADH-Quinone oxidoreductase
164	4633	4651	11	48 h	Ribosomal proteins
165	4690	4734	30	96 h	Ribosomal proteins
182	**5057**	**5100**	**23**	**96 h**	**Actinorhodin biosynthesis**
218c	**5873**	**5886**	**9**	**96 h**	**Undecylprodigiosin biosynthesis**
229	6197	6205	5	96 h	Purine metabolism
232	**6245**	**6288**	**28**	**96 h**	**Coelimycin P1 biosynthesis**
233	6395	6406	7	96 h	β-alanine metabolism
244	6685	6693	5	96 h	Benzoate degradation
248	6968	6982	9	96 h	Inositol metabolism

Function annotation of proteomining clusters were performed by STRING (ver. 11.5). BGCs identified by antiSMASH were highlighted in bold. Clusters with unknown function were omitted.

this cluster was not falsely identified, we obtained the crude methanolic extract of the mycelia at harvested at 96 h and analyzed it by colorimetry and mass spectrometry (Tenconi et al., 2013). The crude extracts exhibited peak absorption at 230 nm and monoisotopic mass of 394.29, consistent with the undecylprodigiosin standard and confirming the activity of the undecylprodigiosin BGC (**Supplementary Figure 2**). On the other hand, the production of coelimycin P1 (PBC No. 232) was only recently documented, due to its BGC expression being controlled by a complex regulatory system (Bednarz et al., 2019).

As such, we believe that the expression of the coelimycin P1 BGC in our study was triggered by the media composition and fermentation conditions used in this study (Gottelt et al., 2010; Bednarz et al., 2019).

Aside from the PBCs identified as BGCs, several PBCs with various functions, determined by STRING, were also upregulated during nutrient starvation. Prominently, members of PBC No. 161 collectively possessed NADH-quinone oxidoreductase functions. The upregulation of quinone oxidoreductases was reasonable due to its correlation with cellular stress and was
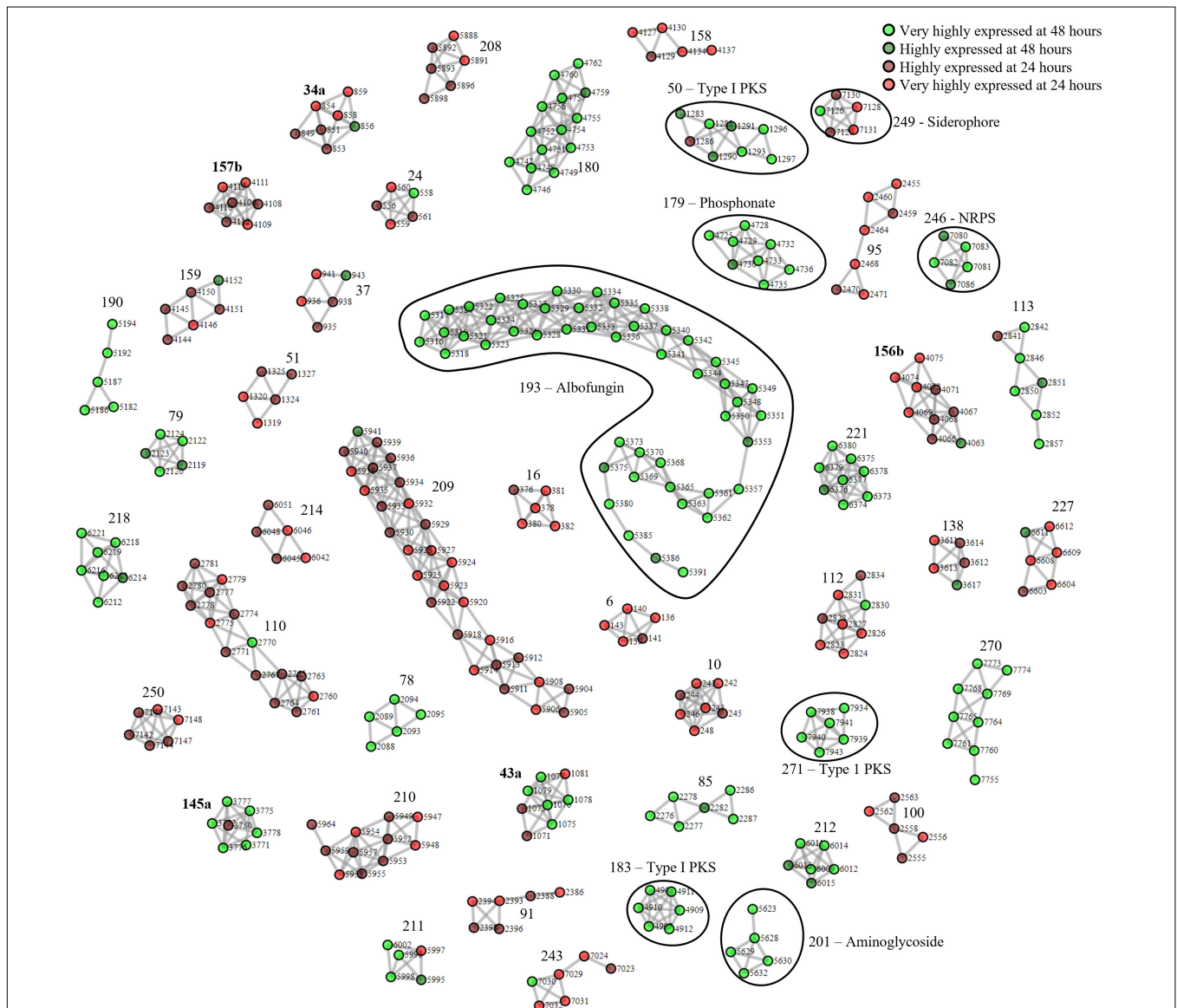

FIGURE 2 | Complete proteomining-based clusters of *S. chrestomyceticus* BCC24770 fermentation in GYM medium with score ≥ 1.5. ● and ● represent proteins that were highly expressed by the mycelia at 48 h (nutrient-starved phase) compared to 24 h (growth phase) at foldchange ≥ 2 and *p*-value ≤ 0.01 (very highly expressed at 48 h), and foldchange ≥ 1.5 and *p*-value ≤ 0.05, respectively, (highly expressed at 48 h). ● and ● represent proteins that were highly expressed by the mycelia at 24 h (growth phase) compared to 48 h (nutrient-starved phase) at foldchange ≥ 2 and *p*-value ≤ 0.01 (very highly expressed at 24 h), and foldchange ≥ 1.5 and *p*-value ≤ 0.05, respectively, (highly expressed at 24 h). Edges were drawn between nodes with an ORF difference ≤ 5. Cluster numbers in bold indicate clusters that previously scored between 1.5 and 3 which scored ≥ 3 when the chimeric clusters were separated.

hypothesized to be involved in radical build-up relief and detoxification of quinones (Pey et al., 2019). Other PBCs that were upregulated during nutrient starvation possessed polysaccharide biosynthesis (PBC No. 7), amino acid transport (PBC No. 69), sensory (PBC No. 89), maleate isomerase (PBC No. 102), purine metabolism (PBC No. 229), β-alanine metabolism (PBC No. 233), benzoate degradation (PBC No. 244) and inositol metabolism (PBC No. 248) functions. On the other hand, numerous PBCs (PBC No. 50, 92, 141b, 164, and 165) containing ribosomal proteins were downregulated during nutrient starvation, consistent with past proteomic studies of *S. coelicolor* A3(2) (Manteca et al., 2006, 2010; Rioseras et al., 2018). This downregulation is likely attributed to nutrient scavenging attempts by recycling ribosomes (Li et al., 2010). Overall, the proteomining clustering method was capable of highlighting active BGCs as well as other active gene clusters.

Streptomyces chrestomyceticus BCC24770 Biosynthetic Gene Cluster Proteomining

Next, we applied our method to *S. chrestomyceticus* BCC24770, which produced a variety of natural products with bioactivity against various nosocomial pathogens (She et al., 2020, 2021). Since the genome of BCC24770 was not fully assembled, being comprised of 70 contigs, we aimed to use this organism to test our proteomining method. Also, unlike *S. coelicolor* A3(2) which exhibited red pigmentation due to accumulation of undecylprodigiosin, the fermentation of BCC24770 did not exhibit detectable cues that indicated production of secondary metabolites. As such the production of albofungin (**Supplementary Figure 3**) was used as the indicator for the nutrient-starved and growth phase, respectively, which were the mycelia harvested from the fermentation in GYM at 48 and 24 h, respectively. The proteomining clustering result yielded 45 PBCs post-decoupling (**Figure 2** and **Supplementary Data 2**). The annotation of detected PBCs was performed by consolidating the function of each ORF in the PBC (**Table 2**). Like the proteomics analysis of *S. coelicolor* A3(2), there was an upregulation in NADH-quinone oxidoreductase activity (PBC No. 212) and a significant downregulation in ribosomal proteins and housekeeping proteins (PBC No. 10, 91, 95, 208–210, and 250).

The PBCs identified were then compared to the genome mining result of antiSMASH. We independently identified PBC No. 193, the largest PBC, as a type II PKS similar to xantholipin BGC. This type II PKS BGC has been characterized to produce albofungin by She et al. Most of the genes in this BGC were detected only during nutrient starvation, except Alb66 (ORF No. 5370. Function: FAD-binding monooxygenase) which was present in both sample groups but still highly upregulated during nutrient starvation (**Supplementary Table 3**). Several transcriptional regulator proteins (ORF No. 5327, 5350, and 5361 corresponding to Alb22, Alb45, and Alb57, respectively) were also only detected during nutrient starvation, suggesting their involvement in eliciting albofungin production.

TABLE 2 | Functional annotation of proteomining clusters of fermentation mycelia *S. chrestomyceticus* BCC24770 in GYM broth at 24 and 48 h.

Cluster	Highly expressed at	Cluster function	Similar BGCs (Top 3)
6	24 h	Iron-sulfur binding proteins	
10	24 h	Cell division-related proteins	
50	**48 h**	**Type I PKS**	**Cycloheximide (27%) 9-Methylstreptimidone (19%)**
91	24 h	Secretion proteins	
95	24 h	Housekeeping proteins	
110	24 h	ATP Synthases	
179	**48 h**	**Phosphonate**	**Dehydrophos (11%)**
183	**48 h**	**NRPS**	**Piericidin A1 (50%)**
193	**48 h**	**Type II PKS (Albofungin)**	
201	**48 h**	**Aminoglycoside***	**Neomycin (85%) Paromomycin (82%)**
208	24 h	Ribosomal proteins/tRNA dimerization	
209	24 h	Ribosomal proteins	
210	24 h	Ribosomal proteins	
212	48 h	NADH-quinone oxidoreductase	
246	**48 h**	**NRPS**	**Mannopeptimycin (81%)**
249	**24 h**	**NRPS (Siderophore)**	**Paenibactin (83%) Streptobactin (70%) Bacillibactin (38%)**
250	24 h	Isochorismate metabolism	
270	48 h	Glycosyl/methyltransferases	
271	**48 h**	**Type I PKS**	**Piericidin A1 (58%)**

BGCs were identified by antiSMASH and is highlighted in bold. * – Only 1 core biosynthetic cluster was differentially expressed between the two sample groups. Clusters with unknown functions were omitted.

Aside from the albofungin BGC, the expression of several smaller orphan BGCs were also detected (**Table 2**). Two type I PKS (PBC No. 50 and 271), 2 NRPS (PBC No. 183 and 246), and 1 phosphonate BGCs (PBC No. 179) were upregulated at the nutrient-starved phase, while PBC No. 249, a putative NRPS BGC that produced siderophores, was the only downregulated BGC. The smaller size of these PBCs, compared to PBC No. 193 or the PBCs identified from *S. coelicolor* A3(2), could be due to the smaller size of the BGCs or inherently lower production of the respective compounds at the time of sampling. Notably, PBC No. 183 and 271 were located near the edge of their respective contigs, as such the smaller size of these 2 PBCs was likely due to an incomplete BGC sequence. Future work could focus on integrating metabolomics in the detection of the products of these orphan BGCs.

To obtain differing fermentation conditions with potentially different secondary metabolite production profiles, we performed fermentation of BCC24770 in TSBY. The fermentation in TSBY was selected due to a stark difference in albofungin production (192–282-fold decrease), which could lead to a different main product of the fermentation (**Supplementary Figure 3**). We then took this opportunity to compare the PBCs identified from the proteomining of BCC24770 fermented in different media to the PBCs identified from the nutrient-starved/growth comparison. Surprisingly, the loss of albofungin production in the fermentation of BCC24770 in TSBY did not result in increased production of other secondary metabolites (**Supplementary Figure 4** and **Supplementary Data 3**). In fact, the PBCs obtained when comparing BCC24770 fermentation in different media were very similar to the nutrient-starved/growth. This suggests a significant downregulation of secondary metabolite biosynthesis of the fermentation in TSBY since some proteins of the albofungin BGC were still expressed at very low concentrations. Some PBCs were also not detected in the proteomining of BCC24770 fermentation in different media, namely a siderophore producing PBC and a putative phosphonate BGC as well as some ribosomal protein PBCs. This could be attributed to a similar trend of downregulation in the mycelia as the culture reaches the stationary phase such as the recycling of ribosomal proteins (Li et al., 2010).

Streptomyces tenebrarius Biosynthetic Gene Cluster Proteomining

Finally, we applied our proteomining clustering algorithm to the apramycin producer, *S. tenebrarius*. Similar to the fermentation of BCC24770, the growth state of *S. tenebrarius* was monitored by the quantity of apramycin in the supernatant (**Supplementary Figure 5**). Apramycin production started between the 12th and 18th hour of fermentation with peak apramycin production between the 18th and 24th hour and plateauing by the 48th hour. The time points 24 and 12 h were appointed as the peak secondary metabolite production phase and the non-producing growth phase, respectively. **Figure 3** displays all the PBCs detected with scores ≥ 3 (**Supplementary Data 4**). Interestingly, only 5 PBCs were expressed, all of which were identified by antiSMASH as BGCs.

Aside from the expression of apramycin BGC, which was monitored throughout the fermentation, the expression of the tobramycin BGC was not unexpected as it has been documented in previous studies of *S. tenebrarius* (Borodina et al., 2005; Hong and Yan, 2012). Interestingly, we also detected the expression of 1 NRPS/Type I PKS hybrid, 1 NRPS and 1 type I PKS BGC which had not been documented in this species before. The closest match for these BGCs were the caerulomycin A (56% similarity), mannopeptimycin (33%) and spectinabilin (81%) BGCs, respectively. We then sought to isolate one of the products of the active BGCs utilizing the ethyl acetate extraction method employed by Zhu et al. (2012) to isolate caerulomycin A. HPLC-MS analysis detected a compound with a mass of 230.09 by HPLC-MS, corresponding to caerulomycin A (**Figure 4**). Here,
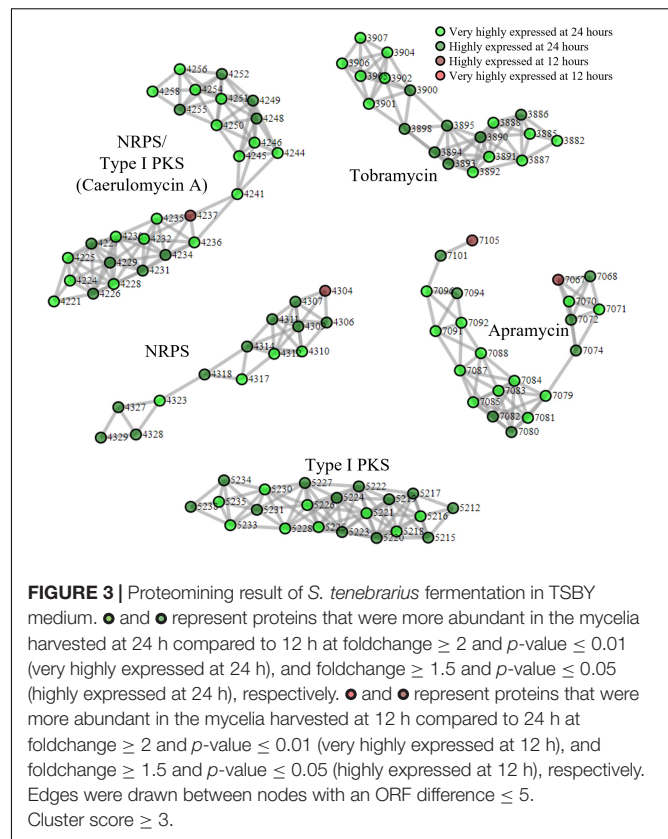

FIGURE 3 | Proteomining result of *S. tenebrarius* fermentation in TSBY medium. ◉ and ◉ represent proteins that were more abundant in the mycelia harvested at 24 h compared to 12 h at foldchange ≥ 2 and *p*-value ≤ 0.01 (very highly expressed at 24 h), and foldchange ≥ 1.5 and *p*-value ≤ 0.05 (highly expressed at 24 h), respectively. ◉ and ◉ represent proteins that were more abundant in the mycelia harvested at 12 h compared to 24 h at foldchange ≥ 2 and *p*-value ≤ 0.01 (very highly expressed at 12 h), and foldchange ≥ 1.5 and *p*-value ≤ 0.05 (highly expressed at 12 h), respectively. Edges were drawn between nodes with an ORF difference ≤ 5. Cluster score ≥ 3.

we demonstrated the use of proteomics to guide secondary metabolite isolation.

Since there was no previous report of the production of caerulomycin A in *S. tenebrarius,* we sought to further confirm the role of the discovered caerulomycin A BGC by gene knockout. We proceeded to knock out ORF 4229–4230 which corresponded to *crmA*, the core biosynthetic protein for caerulomycin A biosynthesis, to probe how this BGC affects secondary metabolite production of *S. tenebrarius*. The loss of *crmA* abolished caerulomycin A production in *S. tenebrarius* (**Figure 4**). Surprisingly, knocking out *crmA* also abolished apramycin production (**Supplementary Figure 5**). We then took this opportunity to test our proteomining clustering method on strain comparison, by comparing *S. tenebrarius* Δ*crmA* and the WT (**Supplementary Figure 6** and **Supplementary Data 5**). Differential expression analysis yielded the same 5 PBCs as the nutrient-starved/growth phase comparison, indicating overall downregulation of secondary metabolite biosynthesis due to the loss of *crmA*. The mechanism behind this cross-BGC regulation is still unclear. Since *crmA* is not a regulatory gene but rather a NRPS that contained a ketosynthase, acyltransferase, acyl-carrier, condensation, adenylation and thiolation domains, this made it unlikely that the overall downregulation of the other BGCs was due to cross-BGC regulation. As such, we hypothesize that the loss of the secondary metabolite BGC expression may be attributed to the loss of caerulomycin A or one of its precursors, which might act as the trigger for secondary
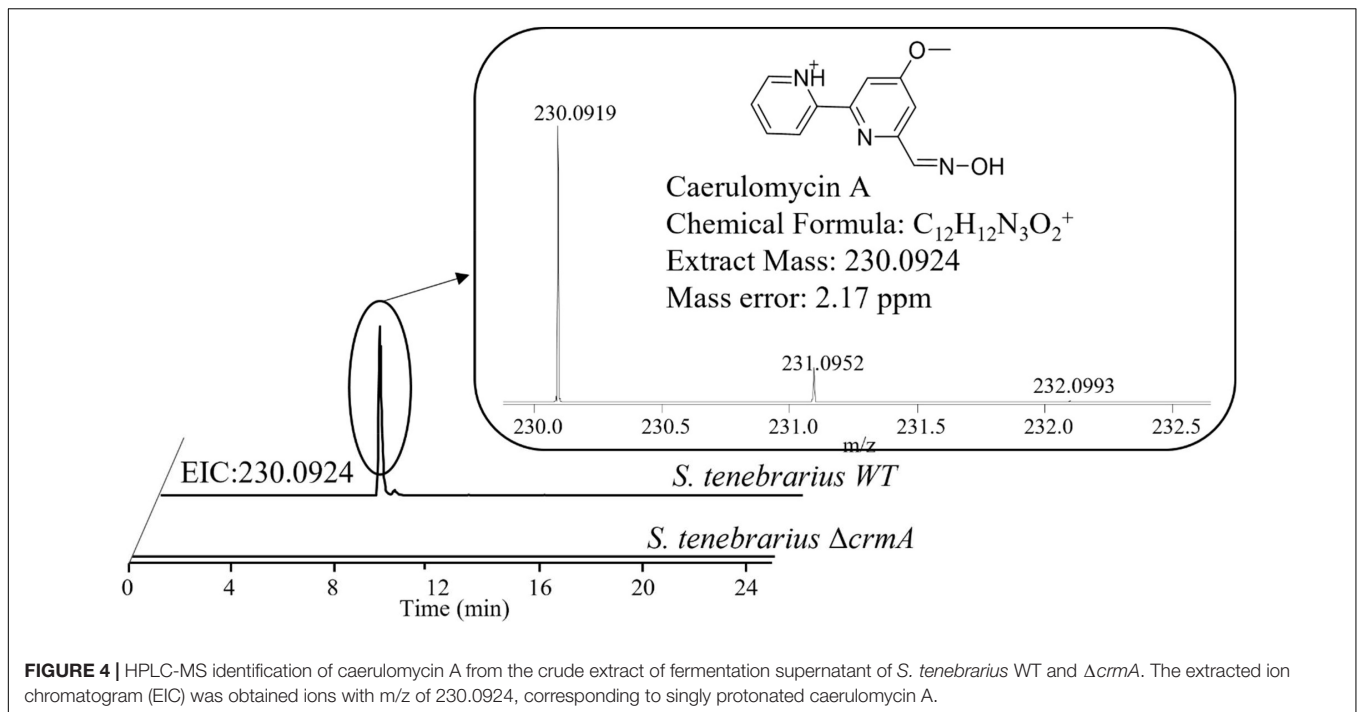
FIGURE 4 | HPLC-MS identification of caerulomycin A from the crude extract of fermentation supernatant of *S. tenebrarius* WT and Δ*crmA*. The extracted ion chromatogram (EIC) was obtained ions with m/z of 230.0924, corresponding to singly protonated caerulomycin A.

metabolite production either by triggering programmed cell death (Tenconi et al., 2018) or through quorum sensing (Koberska et al., 2021).

CONCLUSION

In summary, we developed a proteomining method based on label-free quantitative proteomics of *Streptomyces* sp. We first showed that the different phases of the *Streptomyces* spp. life cycle can be exploited to generate contrasting producing/non-producing states, as demonstrated in a proof of concept in *S. coelicolor* A3(2). Furthermore, the increase in production of secondary metabolite under nutrient starvation conditions is a widespread phenomenon (Hopwood, 2007; Rigali et al., 2008). This approach reduces the need for formulating different fermentation conditions or prior genetic manipulation to produce such contrast, although these other comparisons can also be used.

In addition, recent advancements in LCMS-based mass spectrometry sample preparation techniques and instruments has allowed label-free quantitative proteomics to sufficiently highlight the numerous differentially expressed proteins, while also having the advantages of lower costs, simpler experiments, and the ability to compare across more sample groups, unlike isobaric labeling methods (Du and van Wezel, 2018). Our simple algorithm for clustering differentially expressed proteins by proximity in the genome also facilitated the automated and independent identification of potential BGCs. Compared to the method developed by Gubbens et al. (2014) our algorithm integrates gene proximity and protein differential expression to highlight BGCs more systematically with our scoring system and

directly from label-free quantitative proteomics data, packaged into a simple script. Furthermore, our workflow is simpler, requiring only a single pair of comparison. Overall, our proposed clustering-based proteomining method should be able to detect BGCs with less bias due to non-reliance on co-factors or high molecular weights of proteins (Meluzzi et al., 2008; Bumpus et al., 2009; Chen et al., 2012). Therefore, our proposed workflow, when combined with traditional genome mining, should be helpful in guiding the researcher to isolate novel secondary metabolites produced by *Streptomyces* spp.

In terms of limitations, like other expression-based analysis, proteomining is only applicable when the *Streptomyces* strain is culturable under laboratory conditions and typically relies on nutrient starvation to trigger the production of secondary metabolites (Rigali et al., 2008). In the future, our proteomining clustering algorithm could also be applied in conjunction of the use of engineered media composition or specialized fermentation setup to better simulate the natural habitat of *Streptomyces* sp. which can elicit the expression of silent BGCs (Du and van Wezel, 2018). Another complication is the tendency for *Streptomyces sp.* to activate multiple BGCs to produce distinct secondary metabolites under similar conditions. While the pool of BGC candidates would have been narrower compared to the total BGCs in the genome, it can still be difficult to unambiguously link the BGC to a secondary metabolite of interest based on proteomining alone, though tools such as antiSMASH, PRISM and BLASTP may be called upon to help narrow down the possibilities. In the future, it will be worthwhile to combine time-series metabolomics with proteomics to reveal temporal correlations between individual secondary metabolite production and protein expression changes, to expedite the association of the BGCs to their products with higher confidence and throughput.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: http://www.proteomexchange.org/, PXD032774; http://www.proteomexchange.org/, PXD032775; http://www.proteomexchange.org/, PXD033654. All datasets have been made public.

AUTHOR CONTRIBUTIONS

DL and WS designed and performed the experiments and analyzed the data. QZ performed the gene knockout and caerulomycin A isolation experiments in *S. tenebrarius*. DL wrote the manuscript. HL, YY, and P-YQ supervised the project. All authors revised and edited the manuscript.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2022.913756/full#supplementary-material

REFERENCES

Amos, G. C. A., Awakawa, T., Tuttle, R. N., Letzel, A. C., Kim, M. C., Kudo, Y., et al. (2017). Comparative transcriptomics as a guide to natural product discovery and biosynthetic gene cluster functionality. *Proc. Natl. Acad. Sci. U S A* 114, E11121–E11130. doi: 10.1073/pnas.1714381115

Aziz, R. K., Bartels, D., Best, A., DeJongh, M., Disz, T., Edwards, R. A., et al. (2008). The RAST Server: Rapid annotations using subsystems technology. *BMC Genom.* 9:1–15. doi: 10.1186/1471-2164-9-75

Bednarz, B., Kotowska, M., and Pawlik, K. J. (2019). Multi-level regulation of coelimycin synthesis in Streptomyces coelicolor A3(2). *Appl. Microbiol. Biotechnol.* 103, 6423–6434. doi: 10.1007/s00253-019-09975-w

Bentley, S. D., Chater, K. F., Cerdeño-Tárraga, A. M., Challis, G. L., Thomson, N. R., James, K. D., et al. (2002). Complete genome sequence of the model actinomycete Streptomyces coelicolor A3(2). *Nature* 417, 141–147. doi: 10.1038/417141a

Besemer, J., Lomsadze, A., and Borodovsky, M. (2001). GeneMarkS: A self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. *Nucleic Acids Res.* 29, 2607–2618. doi: 10.1093/nar/29.12.2607

Blin, K., Shaw, S., Kloosterman, A. M., Charlop-Powers, Z., Van Wezel, G. P., Medema, M. H., et al. (2021). AntiSMASH 6.0: Improving cluster detection and comparison capabilities. *Nucleic Acids Res.* 49, W29–W35. doi: 10.1093/nar/gkab335

Borodina, I., Schöller, C., Eliasson, A., and Nielsen, J. (2005). Metabolic network analysis of Streptomyces tenebrarius, a streptomyces species with an active entner-doudoroff pathway. *Appl. Environ. Microbiol.* 71, 2294–2302. doi: 10.1128/AEM.71.5.2294-2302.2005

Bumpus, S. B., Evans, B. S., Thomas, P. M., Ntai, I., and Kelleher, N. L. (2009). A proteomics approach to discovering natural products and their biosynthetic pathways. *Nat. Biotechnol.* 27, 951–956. doi: 10.1038/nbt.1565

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., et al. (2009). BLAST+: Architecture and applications. *BMC Bioinform.* 10:421. doi: 10.1186/1471-2105-10-421

Chambers, M. C., MacLean, B., Burke, R., Amodei, D., Ruderman, D. L., Neumann, S., et al. (2012). A cross-platform toolkit for mass spectrometry and proteomics. *Nat. Biotechnol.* 30, 918–920. doi: 10.1038/nbt.2377

Chen, Y., Unger, M., Ntai, I., McClure, R. A., Albright, J. C., Thomson, R. J., et al. (2013b). Gobichelin A and B: mixed-ligandsiderophores discovered using proteomics. *Med. Chem. Commun.* 4, 233–238. doi: 10.1039/C2MD20232H

Chen, Y., McClure, R. A., Zheng, Y., Thomson, R. J., and Kelleher, N. L. (2013a). Proteomics guided discovery of flavopeptins: Anti-proliferative aldehydes synthesized by a reductase domain-containing non-ribosomal peptide synthetase. *J. Am. Chem. Soc.* 135, 10449–10456. doi: 10.1021/ja4031193

Chen, Y., Ntai, I., Ju, K. S., Unger, M., Zamdborg, L., Robinson, S. J., et al. (2012). A proteomic survey of nonribosomal peptide and polyketide biosynthesis in actinobacteria. *J. Proteome Res.* 11, 85–94. doi: 10.1021/pr2009115

Deutsch, E. W., Mendoza, L., Shteynberg, D., Farrah, T., Lam, H., Sun, Z., et al. (2011). A Guided Tour of the TPP. *Proteomics* 10, 1150–1159. doi: 10.1002/pmic.200900375.A

Du, C., and van Wezel, G. P. (2018). Mining for Microbial Gems: Integrating Proteomics in the Postgenomic Natural Product Discovery Pipeline. *Proteomics* 18, 1–10. doi: 10.1002/pmic.201700332

Evans, B. S., Ntai, I., Chen, Y., Robinson, S. J., and Kelleher, N. L. (2011). Proteomics-based discovery of koranimine, a cyclic imine natural product. *J. Am. Chem. Soc.* 133, 7316–7319. doi: 10.1021/ja2015795

Genilloud, O. (2017). Actinomycetes: Still a source of novel antibiotics. *Nat. Prod. Rep.* 34, 1203–1232. doi: 10.1039/c7np00026j

Genilloud, O. (2018). Mining actinomycetes for novel antibiotics in the omics era: Are we ready to exploit this new paradigm? *Antibiotics* 7:85. doi: 10.3390/antibiotics7040085

Gottelt, M., Kol, S., Gomez-Escribano, J. P., Bibb, M., and Takano, E. (2010). Deletion of a regulatory gene within the cpk gene cluster reveals novel antibacterial activity in Streptomyces coelicolor A3(2). *Microbiology* 156, 2343–2353. doi: 10.1099/mic.0.038281-0

Gubbens, J., Zhu, H., Girard, G., Song, L., Florea, B. I., Aston, P., et al. (2014). Natural product proteomining, a quantitative proteomics platform, allows rapid discovery of biosynthetic gene clusters for different classes of natural products. *Chem. Biol.* 21, 707–718. doi: 10.1016/j.chembiol.2014.03.011

Hao, C., and Lam, H. H. N. (2021). Quantitative Proteomics Reveals UGA-Independent Misincorporation of Selenocysteine throughout the *Escherichia coli* Proteome. *J. Proteome Res.* 20, 212–221. doi: 10.1021/acs.jproteome.0c00352

Hong, W., and Yan, S. (2012). Engineering Streptomyces tenebrarius to synthesize single component of carbamoyl tobramycin. *Lett. Appl. Microbiol.* 55, 33–39. doi: 10.1111/j.1472-765X.2012.03254.x

Hopwood, D. A. (2007). *Streptomyces in Nature and Medicine: The Antibiotic Markers*. Oxford: Oxford Univ. Press.

Kieser, T., Bibb, M. J., Buttner, M. J., and Chater, K. F. (2000). *Practical Streptomyces Genetics. 2nd ed*. England: John Innes Foundation.

Koberska, M., Vesela, L., Vimberg, V., Lenart, J., Vesela, J., Kamenik, Z., et al. (2021). Beyond Self-Resistance: ABCF ATPase LmrC Is a Signal-Transducing Component of an Antibiotic-Driven Signaling Cascade Accelerating the Onset of Lincomycin Biosynthesis. *MBio* 12:e0173121. doi: 10.1128/mBio.01731-21

Li, L., Guo, J., Wen, Y., Chen, Z., Song, Y., and Li, J. (2010). Overexpression of ribosome recycling factor causes increased production of avermectin in Streptomyces avermitilis strains. *J. Ind. Microbiol. Biotechnol.* 37, 673–679. doi: 10.1007/s10295-010-0710-0

Liu, X., Pai, P. J., Zhang, W., Hu, Y., Dong, X., Qian, P. Y., et al. (2016). Proteomic response of methicillin-resistant S. aureus to a synergistic antibacterial drug combination: A novel erythromycin derivative and oxacillin. *Sci. Rep.* 6:19841. doi: 10.1038/srep19841

Ludvigsen, M., and Honoré, B. (2018). *Transcriptomics and Proteomics: Integration?*. Hoboken: Wiley.

Lv, M., Ji, X., Zhao, J., Li, Y., Zhang, C., Su, L., et al. (2016). Characterization of a C3 Deoxygenation Pathway Reveals a Key Branch Point in Aminoglycoside Biosynthesis. *J. Am. Chem. Soc.* 138, 6427–6435. doi: 10.1021/jacs.6b02221

Manteca, A., Jung, H. R., Schwämmle, V., Jensen, O. N., and Sanchez, J. (2010). Quantitative proteome analysis of streptomyces coelicolor nonsporulating liquid cultures demonstrates a complex differentiation process comparable to that occurring in sporulating solid cultures. *J. Proteome Res.* 9, 4801–4811. doi: 10.1021/pr100513p

Manteca, A., Mäder, U., Connolly, B. A., and Sanchez, J. (2006). A proteomic analysis of Streptomyces coelicolor programmed cell death. *Proteomics* 6, 6008–6022. doi: 10.1002/pmic.200600147

McIlwain, S., Mathews, M., Bereman, M. S., Rubel, E. W., MacCoss, M. J., and Noble, W. S. (2012). Estimating relative abundances of proteins from shotgun proteomics data. *BMC Bioinform.* 13:308. doi: 10.1186/1471-2105-13-308

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.

Meier, J. L., Niessen, S., Hoover, H. S., Foley, T. L., Cravatt, B. F., and Burkart, M. D. (2009). An Orthogonal Active Site Identification System (OASIS) for proteomic profiling of natural product biosynthesis. *ACS Chem. Biol.* 4, 948–957. doi: 10.1021/cb9002128

Meier, J. L., Patel, A. D., Niessen, S., Meehan, M., Kersten, R., Yang, J. Y., et al. (2011). Practical 4'-Phosphopantetheine active site discovery from proteomic samples. *J. Proteome Res.* 10, 320–329. doi: 10.1021/pr100953b

Meluzzi, D., Zheng, W. H., Hensler, M., Nizet, V., and Dorrestein, P. C. (2008). Top-down mass spectrometry on low-resolution instruments: Characterization of phosphopantetheinylated carrier domains in polyketide and non-ribosomal biosynthetic pathways. *Bioorganic Med. Chem. Lett.* 18, 3107–3111. doi: 10.1016/j.bmcl.2007.10.104

Moran, M. A., Satinsky, B., Gifford, S. M., Luo, H., Rivers, A., Chan, L. K., et al. (2013). Sizing up metatranscriptomics. *ISME J.* 7, 237–243. doi: 10.1038/ismej.2012.94

Pey, A. L., Megarity, C. F., and Timson, D. J. (2019). NAD(P)H quinone oxidoreductase (NQO1): An enzyme which needs just enough mobility, in just the right places. *Biosci. Rep.* 39, 1–9. doi: 10.1042/BSR20180459

Qu, X., Lei, C., and Liu, W. (2011). Transcriptome mining of active biosynthetic pathways and their associated products in streptomyces flaveolus. *Angew. Chemie – Int. Ed.* 50, 9651–9654. doi: 10.1002/anie.201103085

Rigali, S., Titgemeyer, F., Barends, S., Mulder, S., Thomae, A. W., Hopwood, D. A., et al. (2008). Feast or famine: The global regulator DasR links nutrient stress to antibiotic production by Streptomyces. *EMBO Rep.* 9, 670–675. doi: 10.1038/embor.2008.83

Rioseras, B., Shliaha, P. V., Gorshkov, V., Yagüe, P., López-García, M. T., Gonzalez-Quiñonez, N., et al. (2018). Quantitative proteome and phosphoproteome analyses of streptomyces coelicolor reveal proteins and phosphoproteins modulating differentiation and secondary metabolism. *Mol. Cell. Proteom.* 17, 1591–1611. doi: 10.1074/mcp.RA117.000515

Rossiter, S. E., Fletcher, M. H., and Wuest, W. M. (2017). Natural Products as Platforms To Overcome Antibiotic Resistance. *Chem. Rev.* 117, 12415–12474. doi: 10.1021/acs.chemrev.7b00283

She, W., Ye, W., Cheng, A., Liu, X., Tang, J., Lan, Y., et al. (2021). Discovery, Bioactivity Evaluation, Biosynthetic Gene Cluster Identification, and Heterologous Expression of Novel Albofungin Derivatives. *Front. Microbiol.* 12:635268. doi: 10.3389/fmicb.2021.635268

She, W., Ye, W., Shi, Y., Zhou, L., Zhang, Z., Chen, F., et al. (2020). A novel chresdihydrochalcone from Streptomyces chrestomyceticus exhibiting activity against Gram-positive bacteria. *J. Antibiot.* 73, 429–434. doi: 10.1038/s41429-020-0298-1

Skinnider, M. A., Dejong, C. A., Rees, P. N., Johnston, C. W., Li, H., Webster, A. L. H., et al. (2015). Genomes to natural products PRediction Informatics for Secondary Metabolomes (PRISM). *Nucleic Acids Res.* 43, 9645–9662. doi: 10.1093/nar/gkv1012

Szklarczyk, D., Gable, A. L., Nastou, K. C., Lyon, D., Kirsch, R., Pyysalo, S., et al. (2021). The STRING database in 2021: Customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res.* 49, D605–D612. doi: 10.1093/nar/gkaa1074

Tenconi, E., Guichard, P., Motte, P., Matagne, A., and Rigali, S. (2013). Use of red autofluorescence for monitoring prodiginine biosynthesis. *J. Microbiol. Methods* 93, 138–143. doi: 10.1016/j.mimet.2013.02.012

Tenconi, E., Traxler, M. F., Hoebreck, C., van Wezel, G. P., and Rigali, S. (2018). Production of prodiginines is part of a programmed cell death process in Streptomyces coelicolor. *Front. Microbiol.* 9:1742. doi: 10.3389/fmicb.2018.01742

Zhang, Q., Chi, H. T., Wu, L., Deng, Z., and Yu, Y. (2021). Two Cryptic Self-Resistance Mechanisms in Streptomyces tenebrarius Reveal Insights into the Biosynthesis of Apramycin. *Angew. Chemie – Int. Ed.* 60, 8990–8996. doi: 10.1002/anie.202100687

Zhu, Y., Fu, P., Lin, Q., Zhang, G., Zhang, H., Li, S., et al. (2012). Identification of caerulomycin a gene cluster implicates a tailoring amidohydrolase. *Org. Lett.* 14, 2666–2669. doi: 10.1021/ol300589r