A Rare Thioquinolobactin Siderophore Present in a Bioactive *Pseudomonas* sp. DTU12.1

Pavelas Sazinas , Morten Lindqvist Hansen, May Iren Aune, Marie Højmark Fischer, and Lars Jelsbak *

Department of Biotechnology and Biomedicine, Technical University of Denmark, Kongens Lyngby, Denmark

*Corresponding author: E-mail: lj@bio.dtu.dk.

Accepted: December 2, 2019

Data deposition: This project has been deposited at the Sequencing Read Archive under the accession number PRJNA576879. The complete genome sequence of *Pseudomonas* sp. DTU12.1 has been deposited to Genbank under the accession number CP045254.

Abstract

Many of the soil-dwelling *Pseudomonas* species are known to produce secondary metabolite compounds, which can have antagonistic activity against other microorganisms, including important plant pathogens. It is thus of importance to isolate new strains of *Pseudomonas* and discover novel or rare gene clusters encoding bioactive products. In an effort to accomplish this, we have isolated a bioactive *Pseudomonas* strain DTU12.1 from leaf-covered soil in Denmark. Following genome sequencing with Illumina and Oxford Nanopore technologies, we generated a complete genome sequence with the length of 5,943,629 base pairs. The DTU12.1 strain contained a complete gene cluster for a rare thioquinolobactin siderophore, which was previously described as possessing bioactivity against oomycetes and several fungal species. We placed the DTU12.1 strain within *Pseudomonas gessardii* subgroup of fluorescent pseudomonads, where it formed a distinct clade with other *Pseudomonas* strains, most of which also contained a complete thioquinolobactin gene cluster. Only two other *Pseudomonas* strains were found to contain the gene cluster, though they were present in a different phylogenetic clade and were missing a transcriptional regulator of the whole cluster. We show that having the complete genome sequence and establishing phylogenetic relationships with other strains can enable us to start evaluating the distribution and evolutionary origins of secondary metabolite clusters.

Key words: *Pseudomonas*, whole genome sequencing, secondary metabolite, siderophore, biosynthetic gene cluster, biocontrol.

Introduction

There are multiple species of soil-dwelling fluorescent *Pseudomonas* that are known to be bioactive against other microorganisms, including phytopathogens responsible for various plant root diseases. For example, the virulence of *Dickeya dianthicola*, an important causative agent of blackleg in potatoes, was previously shown to be reduced by several different *Pseudomonas* species, such as *Pseudomonas putida* and *Pseudomonas fluorescens* (Raoul des Essarts et al. 2016). The bioactivity trait of fluorescent pseudomonads is mainly attributed to the production of secondary metabolite compounds, often encoded on large biosynthetic gene clusters (BGCs). These compounds can exert their bioactivity through different modes of action, for example, direct antibiotic activity, induction of systemic resistance in plants or competition for essential nutrients, particularly iron-scavenging by siderophores like pyoverdine and pyochelin (Iavicoli et al. 2003; Deveau et al. 2016). It is thus of extensive agricultural importance to continue isolating new strains of fluorescent pseudomonads and potentially identify novel or rare secondary metabolites that could serve as possible biocontrol agents.

In this study, we investigate the appearance of a rarely reported 8-hydroxy-4-methoxy-2-quinoline thiocarboxylic acid gene cluster in a novel *Pseudomonas* sp. DTU12.1 strain. This bioactive siderophore is also named thioquinolobactin and the biosynthetic pathway for it was first described in *P. fluorescens* strain ATCC17400 (Matthijs et al. 2004; du Moulinet d'Hardemare et al. 2004). The biosynthesis of thioquinolobactin involves tryptophan catabolism via kynurenine pathway and the compound itself is unstable in solution, where it spontaneously hydrolyses to a quinolobactin molecule, whereas retaining siderophore activity (Mossialos et al. 2000). Thioquinolobactin, though not quinolobactin, was shown to be bioactive against an oomycete phytopathogen

© The Author(s) 2019. Published by Oxford University Press on behalf of the Society for Molecular Biology and Evolution.
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

Pythium debaryanum and several phytopathogenic fungal species under iron-limiting conditions (Matthijs et al. 2007). Despite the potential biocontrol applications of thioquinolobactin-producing *Pseudomonas* strains, the distribution and frequency of the thioquinolobactin BGC in the *Pseudomonas* genus remain unknown. Here, we describe the novel *Pseudomonas* sp. DTU12.1 strain and address the question of the phylogenetic distribution and possible evolutionary history of the thioquinolobactin gene cluster.

Materials and Methods

For the isolation of bioactive *Pseudomonas* species, collected soil sample was initially diluted in 0.9% sodium chloride solution and plated on selective *Pseudomonas* agar, containing 13 μg/ml chloramphenicol. The agar plates were incubated for 48 h at room temperature and fluorescent colonies were subsequently picked. 16S rDNA Sanger sequencing was performed using universal 16S rDNA gene primers 8F (5′-AGAGTTTGATCCTGGCTCAG-3′) and 1492R (5′-GGTTACCTTGTTACGACTT-3′). For Illumina sequencing, whole genome DNA was isolated with Wizard Genomic DNA Purification kit (Promega) and genomic library preparation was done using modified (half volume for each reagent) Kapa Hyper Plus Library Prep kit (Roche Molecular Systems). The Illumina sequencing was carried out on Illumina MiSeq (300 cycles). Trimming of low quality Illumina reads (<Q20) was performed with trimmomatic v0.38. For Oxford Nanopore, whole genome DNA was extracted with PureLink Genomic DNA kit (Thermo Fisher Scientific) and genomic library was prepared using a Ligation kit (Oxford Nanopore). The Oxford Nanopore sequencing was carried out on the MinION device using Flow Cell FLO-MIN106.

A hybrid genome assembly was done with Unicycler v0.4.6 and genome annotation was performed by Prokaryotic Genome Annotation Pipeline (PGAP) (Tatusova et al. 2016; Wick et al. 2017). Putative BGCs were identified by a combination of the AntiSMASH v3.0 database and manual curation (Weber et al. 2015). A Neighbor-Joining phylogenetic tree was generated in MEGA X v10.1 software, where nucleotide sequences were aligned with the MUSCLE algorithm, followed by constructing a bootstrap consensus tree (1,000 replicates) (Kumar et al. 2008). BLASTP was used to identify homologous genes above specified cut-offs (75% query coverage, 75% identity). MultiGeneBlast v1.1.14 software was used to perform a gene cluster architecture search against BLASTP hits using the same cut-off values (Medema et al. 2013). As the first step in species prediction, the method by Mulet et al. (2010) was used, where BlastN comparison of 16S rDNA, *rpoD*, *rpoB*, and *gyrB* genes was done across all *Pseudomonas* species (Mulet et al. 2010). As the second step in species prediction, JSpeciesWS online tool was used to perform average nucleotide identity analysis (ANIb) on selected genomes (Richter et al. 2016).

Results and Discussion

In an effort to isolate bioactive *Pseudomonas* species and characterize any rare or even novel secondary metabolites they might be producing, several soil samples were collected from different sites in Denmark. Following the plating of collected samples on selective *Pseudomonas* agar, multiple fluorescent colonies were picked and 16S rDNA sequencing was carried out to check if they belonged to the *Pseudomonas* genus. One of the confirmed isolated *Pseudomonas* strains was designated as DTU12.1. Its bioactivity was tested against five plant pathogens and a growth inhibition zone was observed for four of the pathogens (*Xanthomonas campestris* and three *Pectobacterium carotovora* subspecies), indicating that DTU12.1 is a bioactive strain (our unpublished data). Therefore, it was deemed a good candidate for further genomic analysis.

We sequenced the genome of DTU12.1 using both Illumina and Oxford Nanopore technologies. A total of 5,218,438 paired-end (2 × 150 bp) reads were produced with Illumina. With Oxford Nanopore, 44,049 reads were generated above the Q10 quality filter. All sequencing reads have been deposited into the Sequencing Read Archive (PRJNA576879). The mean length of Nanopore reads was 7,000 bp and N50 value for the read length was 14,561 bp. Following a hybrid genome assembly, a circular complete chromosome was generated, with the length of 5,943,629 bp, G + C content of 60.68% and read depth of 120.9× (see supplementary table 1, Supplementary Material online). The complete genome sequence was submitted to Genbank under accession number CP045254. As determined through annotation by PGAP, the number of genes was 5,454, which included 5,245 protein-coding genes, 19 rRNAs, 69 tRNAs, 4 noncoding RNAs, and 117 pseudogenes.

In order to start inferring the phylogenetic relationships of *Pseudomonas* sp. DTU12.1, an *rpoD*-based phylogenetic analysis was done on all representative *Pseudomonas* species (Hesse et al. 2018). DTU12.1 was placed within a *Pseudomonas gessardii* subgroup of fluorescent pseudomonads (see fig. 1). We then set out to investigate the presence of secondary metabolite clusters in the genome of DTU12.1. On the basis of the output from the AntiSMASH database, 11 putative BGCs were identified (see supplementary table 2, Supplementary Material online). One of these was a *qbs* gene cluster coding for a bioactive thioquinolobactin siderophore (Matthijs et al. 2004). The DTU12.1 strain contained the complete cluster of 15 genes, with locus tags GDV60_15535-GDV60_15605. As the thioquinolobactin was previously found to be bioactive against oomycetes and some fungal species, it is possible that it also contributes to bioactivity against bacterial phytopathogens tested with DTU12.1.

To determine which other *Pseudomonas* strains contained the thioquinolobactin BGC, we carried out a BLASTP search of individual genes, followed by a gene architecture analysis with

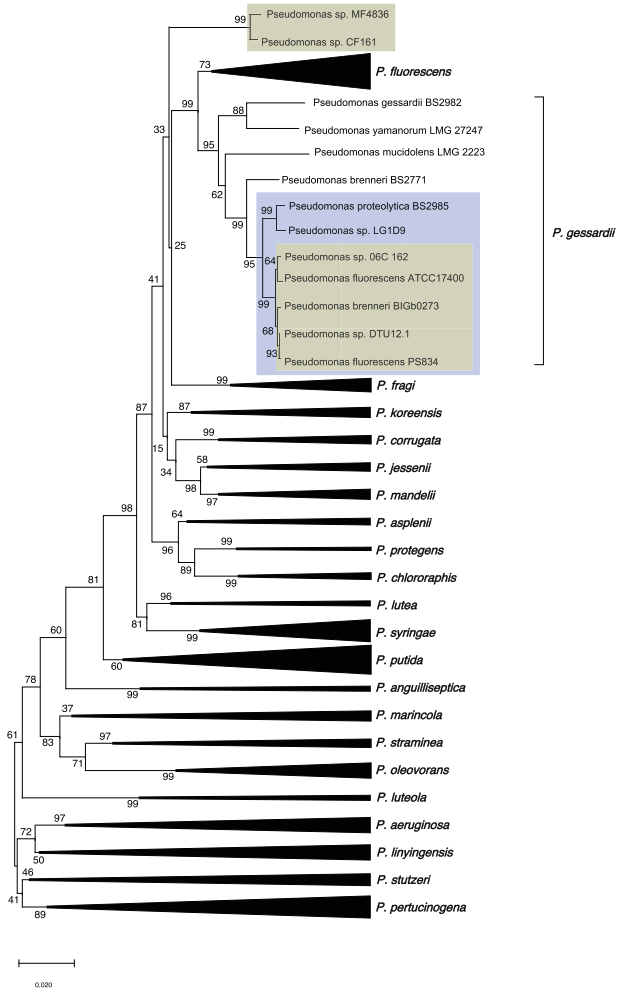

Fig. 1.—An *rpoD*-based Neighbor-Joining tree for 164 type strains of the *Pseudomonas* genus and eight other specifically selected *Pseudomonas* strains, including the newly isolated *Pseudomonas* sp. DTU12.1. MEGA X v10.1 was used to align the *rpoD* gene nucleotide sequences with the MUSCLE algorithm, followed by generating a bootstrap consensus tree (1,000 replicates) via the Neighbor-Joining method. The bootstrap percentage values are depicted next to each branching point. Groups and subgroups of *Pseudomonas* are compressed in clades and named based on Hesse et al. (2018), where the size of the clade is proportional to the number of species in it. The *Pseudomonas* strains that are the same species as *Pseudomonas* sp. DTU12.1 (as based on the ANIb analysis) are placed within a blue box, while the strains containing the thioquinolobactin gene cluster are present inside green boxes. The scale bar represents the number of nucleotide substitutions per site.

MultiGeneBlast above specified cut-offs (see fig. 2). There were only four other strains (*P. fluorescens* ATCC17400, *P. fluorescens* PS834, *Pseudomonas brenneri* BIGb0273, and *Pseudomonas* sp. 06C 126) where the complete gene cluster was present. In the 06C 126 strain the thioquinolobactin gene cluster was previously shown to contribute to the growth inhibition of oomycetes and *Pseudomonas aeruginosa* isolates from cystic fibrosis patients (Chatterjee et al. 2017; Wagner et al. 2018). There were also two other strains, *Pseudomonas* sp. CF161 and MF4836, where the thioquinolobactin BGC was missing *orfP* and *qbsA* genes, the latter encoding a transcriptional regulator (see fig. 2).

We wanted to understand the relationship between *Pseudomonas* sp. DTU12.1 and other strains containing the thioquinolobactin gene cluster as well as determine the degree of similarity between DTU12.1 and its most related strains. A taxonomic assignment of DTU12.1 using two approaches was therefore carried out. We first did a BlastN analysis on four housekeeping genes (16S rDNA, *rpoD*, *rpoB*, and *gyrB*) of DTU12.1 against all *Pseudomonas* genomes (Mulet et al. 2010). The single top hit with an identity score of 99.73% was *Pseudomonas* sp. LG1D9. This genome, along with all the identified strains containing the thioquinolobactin gene cluster as well as three phylogenetically adjacent type strains (*Pseudomonas proteolytica* BS2985, *P. gessardii* BS2985, and *P. brenneri* BS2771) were then used in the ANIb analysis. All of the strains containing the complete thioquinolobactin gene cluster had high ANIb value of >98% among each other, suggesting that they all belong to the same species (see supplementary table 3, Supplementary Material online). This is despite the discrepancies in the naming, as some of these were previously assigned as different species, for example, *P. brenneri* BIGb2073 and *P. fluorescens* ATCC17400. The exact ANIb values for the closest relatives of DTU12.1, *P. brenneri* BIGb2073, *Pseudomonas* sp. LG1D9, *P. proteolytica* BS2985, *P. fluorescens* PS834, *P. fluorescens* ATCC17400 and *Pseudomonas* sp. 06C 126, were 98.65%, 98.63%, 98.48%, 98.33%, 98.29%, and 98.26%, respectively, signifying over 75,000 nucleotide differences between these strains and DTU12.1. While both *Pseudomonas* sp. LG1D9 and *P. proteolytica* BS2985 were part of the same species based on ANIb score, they did not contain the thioquinolobactin cluster. *Pseudomonas* sp. CF161 and MF4836, both possessing an incomplete thioquinolobactin cluster, had high ANIb score between each other, implying that they both belong to the same species, though separate from *Pseudomonas* sp. DTU12.1 and others.

We then revisited the *rpoD*-based phylogenetic tree by including all the strains containing the thioquinolobactin gene cluster as well as *Pseudomonas* sp. LG1D9. The four strains with the complete set of 15 genes appeared within the same phylogenetic clade as DTU12.1 (see fig. 1). *Pseudomonas* sp. LG1D9 and *P. proteolytica* BS2985, both without the thioquinolobactin gene cluster, also clustered close to DTU12.1 and others, confirming that all species with the high ANIb score (>95%) phylogenetically group together. The two other strains, CF161 and MF4836, both missing *qbsA* and *orfP* genes, formed their own distinct clade, which was placed outside *P. gessardii* subgroup. It was previously shown that a *qbsA* mutant of *P. fluorescens* ATCC17400 produces thioquinolobactin at a much reduced level (Matthijs et al. 2004). It was therefore hypothesized that *qbsA* encodes a

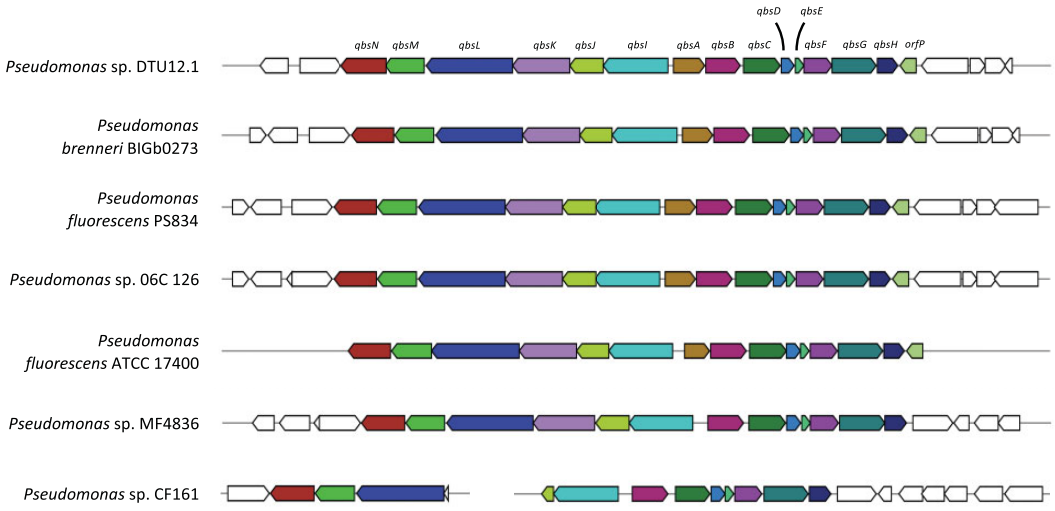

Fig. 2.—A gene architecture analysis of the thioquinolobactin gene cluster, as inferred with the MultiGeneBlast software. The genes in the *qbs* cluster are named based on the original annotation in *Pseudomonas fluorescens* ATCC17400. The same colors represent homologous genes in different species. In *Pseudomonas* sp. CF161 the gene cluster is split between two contigs.

transcriptional activator of the whole cluster, which is additionally regulated by Fur. It is possible that the acquisition of *qbsA* was an important evolutionary step toward the high expression of thioquinolobactin gene cluster in DTU12.1 and closely related strains.

We have presented in this study that the thioquinolobactin gene cluster is observed very rarely within the *Pseudomonas* genus. The polyphyletic distribution of the gene cluster is akin to previously reported studies on BGCs. For example, the siderophore pyoverdine gene cluster was shown to exist disparately in fluorescent pseudomonads as well as being present in more distantly related species like *P. aeruginosa* (Hesse et al. 2018). It has also been previously described that a 2, 4-diacetylphloroglucinol (DAPG) gene cluster is present in separate phylogenetic groups of *Pseudomonas* species, suggesting multiple acquisition events rather than vertical inheritance (Almario et al. 2017). It is possible that there was a horizontal gene transfer event where the ancestor of DTU12.1 and related strains obtained the incomplete gene cluster, followed by the acquisition of the transcriptional activator *qbsA*, which led to the high expression of the cluster.

We have shown that the screening of soil-dwelling *Pseudomonas* species for bioactivity against various plant pathogens can lead to the identification of rare secondary metabolite gene clusters. The whole genome sequencing and genomic characterization of fluorescent pseudomonads is a necessary approach in order to start understanding how secondary metabolite gene clusters evolve and disseminate within the *Pseudomonas* genus. A better understanding of the phylogenetic distribution of the BGCs that encode bioactive metabolites can help to design more efficient sampling and screening strategies for strains with bioactivity against plant pathogens.

Supplementary Material

Supplementary data are available at *Genome Biology and Evolution* online.

Acknowledgments

This study was supported by funding from the Danish National Research Foundation (DNRF137) for the Center for Microbial Secondary Metabolites.

Literature Cited

Almario J et al. 2017. Distribution of 2, 4-diacetylphloroglucinol biosynthetic genes among the *Pseudomonas* spp. Reveals unexpected polyphyletism. Front Microbiol. 8:1218.

Chatterjee P, et al. 2017. Environmental pseudomonads inhibit cystic fibrosis patient-derived *Pseudomonas aeruginosa*. Appl Environ Microbiol. 83(2):e02701–16.

Deveau A, et al. 2016. Role of secondary metabolites in the interaction between *Pseudomonas fluorescens* and soil microorganisms under iron-limited conditions. FEMS Microbiol Ecol. 92(8):fiw107.

du Moulinet d'Hardemare A, Serratrice G, Pierre J-L. 2004. Synthesis and iron-binding properties of quinolobactin, a siderophore from a pyoverdine-deficient *Pseudomonas fluorescens*. Biometals 17:691–697.

Hesse C, et al. 2018. Genome-based evolutionary history of *Pseudomonas* spp. Environ Microbiol. 20(6):2142–2159.

Iavicoli A, Boutet E, Buchala A, Métraux J-P. 2003. Induced systemic resistance in *Arabidopsis thaliana* in response to root inoculation with *Pseudomonas fluorescens* CHA0. MPMI 16(10):851–858.

Kumar S, Nei M, Dudley J, Tamura K. 2008. MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. Brief Bioinform. 9(4):299.

Matthijs S, et al. 2004. The *Pseudomonas* siderophore quinolobactin is synthesized from xanthurenic acid, an intermediate of the kynurenine pathway. Mol Microbiol. 52(2):371–384.
Matthijs S, et al. 2007. Thioquinolobactin, a Pseudomonas siderophore with antifungal and anti-*Pythium* activity. Environ Microbiol. 9(2):425–434.
Medema MH, Takano E, Breitling R. 2013. Detecting sequence homology at the gene cluster level with MultiGeneBlast. Mol Biol Evol. 30(5):1218–1223.
Mossialos D, et al. 2000. Quinolobactin, a new siderophore of *Pseudomonas fluorescens* ATCC 17400, the production of which is repressed by the cognate pyoverdine. Appl Environ Microbiol. 66(2):487–492.
Mulet M, Lalucat J, Garcia-Valdes E. 2010. DNA sequence-based analysis of the *Pseudomonas* species. Environ Microbiol. 12:1513–1530.
Raoul des Essarts Y, et al. 2016. Biocontrol of the potato blackleg and soft rot diseases caused by *Dickeya dianthicola*. Appl Environ Microbiol. 82(1):268–278.
Richter M, Rossello-Mora R, Oliver Glockner F, Peplies J. 2016. JSpeciesWS: a web server for prokaryotic species circumscription based on pairwise genome comparison. Bioinformatics 32(6):929–931.
Tatusova T, et al. 2016. NCBI prokaryotic genome annotation pipeline. Nucleic Acids Res. 44(14):6614–6624.
Wagner A, Norris S, Chatterjee P, Morris PF, Wildschutte H. 2018. Aquatic pseudomonads inhibit oomycete plant pathogens of glycine max. Front Microbiol. 9:1007.
Weber T, et al. 2015. antiSMASH 3.0—a comprehensive resource for the genome mining of biosynthetic gene clusters. Nucleic Acids Res. 43(W1):W237–43.
Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput Biol. 13(6):e1005595.

Associate editor: Howard Ochman